Lecture 2: The Volatility Surface

Jim Gatheral, Merrill Lynch

Case Studies in Financial Modelling Course Notes,
Courant Institute of Mathematical Sciences,
Fall Term, 2004

4 Getting Implied Volatility from Local Volatilities

4.1 Model Calibration

For a model to be useful in practice, it needs to return (at least approximately) the current market prices of European options. That implies that we need to fit the parameters of our model (whether stochastic or local volatility model) to market implied volatilities. It is clearly easier to calibrate a model if we have a fast and accurate method for computing the prices of European options as a function of the model parameters. In the case of stochastic volatility, this consideration clearly favors models such as Heston which have such a solution; Mikhailov and Nögel (July 2003) for example explain how to calibrate the Heston model to market data.

In the case of local volatility models, numerical methods are usually required to compute European option prices and that is one of the potential problems associated with their implementation. Brigo and Mercurio (2002) circumvent this problem by parameterizing the local volatility in such a way that the prices of European options are known in closed-form as superpositions of Black-Scholes-like solutions.

Yet again, we could work with the European option prices directly in a trinomial tree framework as in Derman, Kani, and Chriss (1996) or we could maximize relative entropy (of missing information) as in Avellaneda, Friedman, Holmes, and Samperi (1997). These methods are non-parametric (assuming actual option prices are used, not interpolated or extrapolated values); they may fail because of noise in the prices and the bid/offer spread.

Finally, we could parameterize the risk-neutral distributions as in Rubinstein (1998) or parameterize the implied volatility surface directly as in Shimko (1993). Although these approaches look straightforward given that we know from Section 2.3 how to get local volatility in terms of implied volatility, these approaches are very difficult to implement in practice. The problem is that we don't have a complete implied volatility surface, we only have a few bids and offers per expiration. To apply a parametric method, we need to interpolate and extrapolate the known implied volatilities. It is very hard to do this without introducing arbitrage. The arbitrages to avoid are roughly speaking, negative vertical spreads, negative butterflies and negative calendar spreads (where the latter are carefully defined).

In what follows, we will concentrate on the implied volatility structure of stochastic volatility models so we won't have to worry about the possibility of arbitrage which is excluded from the outset.

First, we derive an expression for implied volatility in terms of local volatilities. In principle, this should allow us to investigate the shape of the implied volatility surface for any local volatility or stochastic volatility model because we know from Section 2.5 how to express local variance as an expectation of instantaneous variance in a stochastic volatility model.

4.2 Understanding Implied Volatility[1]

In Section 2.3, we derived an expression for local volatility in terms of implied volatility. An obvious direct approach might be to invert that expression and express implied volatility in terms of local volatility. However, this kind of direct attack on the problem doesn't yield any easy results (at least not to me).

Instead, by extending the work of Blacher (1998), we derive a general path-integral representation of Black-Scholes implied variance. We start by assuming that the stock price S_t satisfies the SDE

$$\frac{dS_t}{S_t} = \mu_t dt + \sigma_t \, dZ_t$$

where the volatility σ_t may be random.

For fixed K and T, define the Black-Scholes gamma

$$\Gamma_{BS}(S_t, \bar{\sigma}(t)) := \frac{\partial^2}{\partial S_t^2} C_{BS}(S_t, K, \bar{\sigma}(t), T - t)$$

and further define the "Black-Scholes forward implied variance" function

$$v_{K,T}\,(t) = \frac{\mathbb{E}\left[\sigma_t^2 \, S_t^2 \, \Gamma_{BS}(S_t, \bar{\sigma}(t)) \,|\mathcal{F}_0\right]}{\mathbb{E}\left[S_t^2 \, \Gamma_{BS}(S_t, \bar{\sigma}(t)) \,|\mathcal{F}_0\right]} \tag{24}$$

where

$$\bar{\sigma}^2(t) := \frac{1}{T-t} \int_t^T v_{K,T}(u) \, du \tag{25}$$

[1]The rearrangement and clarification of the material in this section owes much to Roger Lee (see Lee (2002)).

Path-by-path, for any suitably smooth function $f(S_t, t)$ of the random stock price S_t and for any given realization $\{\sigma_t\}$ of the volatility process, the difference between the initial value and the final value of the function $f(S_t, t)$ is obtained by anti-differentiation. Then, applying Itô's Lemma, we get

$$\begin{aligned} f(S_T, T) - f(S_0, 0) &= \int_0^T df \\ &= \int_0^T \left\{ \frac{\partial f}{\partial S_t} \, dS_t + \frac{\partial f}{\partial t} \, dt + \frac{\sigma_t^2}{2} \, S_t^2 \, \frac{\partial^2 f}{\partial S_t^2} \, dt \right\} \end{aligned} \tag{26}$$

Under the usual assumptions, the non-discounted value $C(S_0, K, T)$ of a call option is given by the expectation of the final payoff under the risk-neutral measure. Then, applying (26), we obtain:

$$\begin{aligned} C(S_0, K, T) &= \mathbb{E}\left[(S_T - K)^+ \, | \mathcal{F}_0 \right] \\ &= \mathbb{E}\left[C_{BS}(S_T, K, \bar{\sigma}(T), 0) \, | \mathcal{F}_0 \right] \\ &= C_{BS}(S_0, K, \bar{\sigma}(0), T) \\ &\quad + \mathbb{E}\left[\int_0^T \left\{ \frac{\partial C_{BS}}{\partial S_t} dS_t + \frac{\partial C_{BS}}{\partial t} dt + \frac{1}{2} \sigma_t^2 S_t^2 \frac{\partial^2 C_{BS}}{\partial S_t^2} dt \right\} \right| \left. \mathcal{F}_0 \right] \end{aligned}$$

Now of course $C_{BS}(S_t, K, \bar{\sigma}(t), T - t)$ must satisfy the Black-Scholes equation (assuming zero interest rates and dividends) and from the definition of $\bar{\sigma}(t)$, we obtain:

$$\frac{\partial C_{BS}}{\partial t} = -\frac{1}{2} v_{K,T}(t) \; S_t^2 \, \frac{\partial^2 C_{BS}}{\partial S_t^2}$$

Using this equation to substitute for the time derivative $\frac{\partial C_{BS}}{\partial t}$, we obtain:

$$\begin{aligned} C(S_0, K, T) &= C_{BS}(S_0, K, \bar{\sigma}(0), T) \\ &\quad + \mathbb{E}\left[\int_0^T \left\{ \frac{\partial C_{BS}}{\partial S_t} dS_t + \frac{1}{2} \left\{ \sigma_t^2 - v_{K,T}(t) \right\} S_t^2 \frac{\partial^2 C_{BS}}{\partial S_t^2} dt \right\} \right| \left. \mathcal{F}_0 \right] \\ &= C_{BS}(S_0, K, \bar{\sigma}(0), T) \\ &\quad + \mathbb{E}\left[\int_0^T \frac{1}{2} \left\{ \sigma_t^2 - v_{K,T}(t) \right\} S_t^2 \frac{\partial^2 C_{BS}}{\partial S_t^2} dt \right| \left. \mathcal{F}_0 \right] \end{aligned} \tag{27}$$

where the second equality uses the fact that S_t is a martingale.

From the definition (25) of $v_{K,T}(t)$, we have that

$$\mathbb{E}\left[S_t^2\,\Gamma_{BS}(S_t,\bar{\sigma}(t))\,|\mathcal{F}_0\right]v_{K,T}\,(t)=\mathbb{E}\left[\sigma_t^2\,S_t^2\,\Gamma_{BS}(S_t,\bar{\sigma}(t))\,|\mathcal{F}_0\right]$$

so the second term in equation (27) vanishes and from the definition of implied volatility, $\bar{\sigma}(0)$ is the Black-Scholes implied volatility at time 0 of the option with strike K and expiration T (*i.e.* the Black-Scholes formula must give the market price of the option).

Explicitly

$$\sigma_{BS}(K,T)^2=\bar{\sigma}(0)^2=\frac{1}{T}\int_0^T\frac{\mathbb{E}\left[\sigma_t^2S_t^2\Gamma_{BS}(S_t)\,|\mathcal{F}_0\right]}{\mathbb{E}\left[S_t^2\Gamma_{BS}(S_t)\,|\mathcal{F}_0\right]}dt \tag{28}$$

Equation (28) expresses implied variance as the time-integral of expected instantaneous variance σ_t^2 under some probability measure.

More precisely, following Lee (2002), we may write

$$\sigma_{BS}(K,T)^2=\bar{\sigma}(0)^2=\frac{1}{T}\int_0^T\mathbb{E}^{G_t}[\sigma_t^2]\,dt \tag{29}$$

thus interpreting the definition (24) of $v(t)$ as the expectation of σ_t^2 with respect to the probability measure $\mathbb{G}_t$ defined, relative to the pricing measure $\mathbb{P}$, by the Radon-Nikodym derivative

$$\frac{d\mathbb{G}_t}{d\mathbb{P}}:=\frac{S_t^2\,\Gamma_{BS}(S_t,\bar{\sigma}(t))}{\mathbb{E}\left[S_t^2\,\Gamma_{BS}(S_t,\bar{\sigma}(t))\,|\mathcal{F}_0\right]}$$

Note in passing that equations (24) and (28) are implicit because the gamma $\Gamma_{BS}(S_t)$ of the option depends on all the forward implied variances $v_{K,T}\,(t)$.

Special Case (Black-Scholes)

Suppose $\sigma_t=\sigma(t)$, a function of t only. Then

$$v_{K,T}\,(t)=\frac{\mathbb{E}\left[\sigma(t)^2S_t^2\Gamma_{BS}(S_t)\,|\mathcal{F}_0\right]}{\mathbb{E}\left[S_t^2\Gamma_{BS}(S_t)\,|\mathcal{F}_0\right]}=\sigma(t)^2$$

The forward implied variance $v_{K,T}(t)$ and the forward variance $\sigma(t)^2$ coincide. As expected, $v_{K,T}(t)$ has no dependence on the strike K or the option expiration T.

Interpretation

In order to get better intuition for equation (24), first recall how to compute a risk-neutral expectation:

$$\mathbb{E}^P\left[f\left(S_t\right)\right] = \int dS_t\, p\left(S_t, t; S_0\right) f\left(S_t\right)$$

We get the risk-neutral pdf of the stock price at time t by taking the second derivative of the market price of European options with respect to strike price.

$$p\left(S_t, t; S_0\right) = \left.\frac{\partial^2 C\left(S_0, K, t\right)}{\partial K^2}\right|_{K=S_t}$$

Then from equation (29) we have

$$\begin{aligned} v_{K,T}\left(t\right) &= \mathbb{E}^{G_t}\left[\sigma_t^2\right] \\ &= \mathbb{E}^P\left[\sigma_t^2 \frac{d\mathbb{G}_t}{d\mathbb{P}}\right] \\ &= \int dS_t\, q\left(S_t; S_0, K, T\right) \mathbb{E}^P\left[\sigma_t^2 | S_t\right] \\ &= \int dS_t\, q\left(S_t; S_0, K, T\right) v_L(S_t, t) \end{aligned} \tag{30}$$

where we further define

$$q\left(S_t, t; S_0, K, T\right) := \frac{p\left(S_t, t; S_0\right)\ S_t^2\, \Gamma_{BS}(S_t)}{\mathbb{E}\left[S_t^2\, \Gamma_{BS}(S_t)\,|\mathcal{F}_0\right]}$$

and $v_L(S_t, t) = \mathbb{E}^P\left[\sigma_t^2 | S_t\right]$ is the local variance.

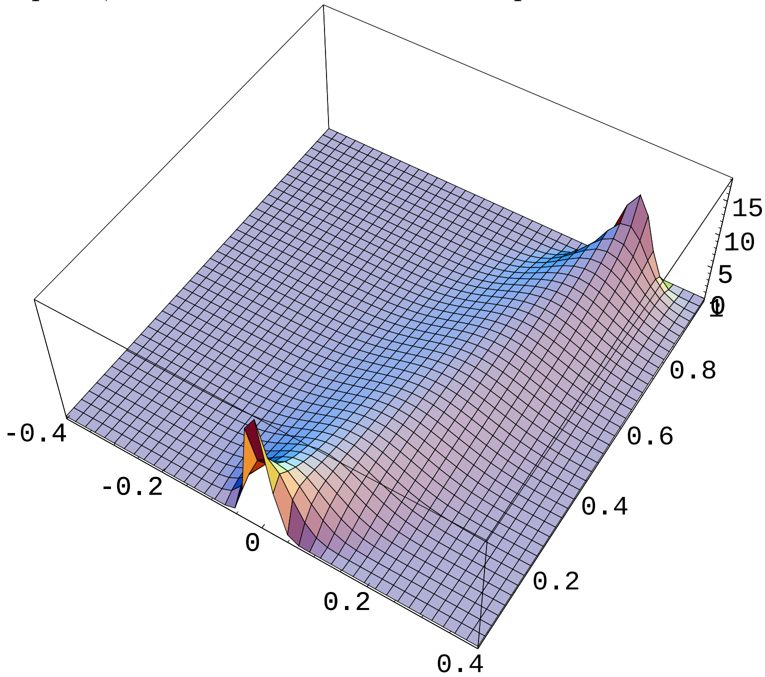
We see that $q\left(S_t, t; S_0, K, T\right)$ looks like a Brownian Bridge density for the stock price: $p\left(S_t, t; S_0\right)$ has a delta-function peak at S_0 at time 0 and $\Gamma_{BS}(S_t)$ has a delta-function peak at K at expiration T.

For convenience in what follows, we now rewrite equation (30) in terms of $x_t \equiv \log\left(S_t/S_0\right)$:

$$v_{K,T}\left(t\right) = \int dx_t\, q\left(x_t, t; x_T, T\right) v_L(x_t, t) \tag{31}$$

Figure 1 shows how $q\left(x_t, t; x_T, T\right)$ looks in the case of a 1 year European option struck at 1.3 with a flat 20% volatility. We see that $q\left(x_t, t; x_T, T\right)$

Figure 1: Graph of the pdf of x_t conditional on $x_T = \log(K)$ for a 1 year European option, strike 1.3 with current stock price = 1 and 20% volatility.



peaks on a line (which we will denote by $\tilde{x}_t$) joining the stock price today with the strike price at expiration. Moreover, the density looks roughly symmetric around the peak. This suggests an expansion around the peak $\tilde{x}_t$ (at which the derivative of $q\left(x_t, t; x_t, T\right)$ with respect to x_t is zero). Then we write:

$$q\left(x_t, t; x_T, T\right) \approx q(\tilde{x}_t, t; x_T, T) + \frac{1}{2}\left(x_t - \tilde{x}_t\right)^2 \left.\frac{\partial^2 q}{\partial x_t^2}\right|_{x_t = \tilde{x}_t} \tag{32}$$

In practice, the local variance $v_L(x_t, t)$ is typically not so far from linear in x_t in the region where $q\left(x_t, t; x_T, T\right)$ is significant so we may further write

$$v_L(x_t, t) \approx v_L(\tilde{x}_t, t) + \left(x_t - \tilde{x}_t\right) \left.\frac{\partial v_L}{\partial x_t}\right|_{x_t = \tilde{x}_t} \tag{33}$$

Substituting (32) and (33) into the integrand in equation (31) gives

$$v_{K,T}\left(t\right) \approx v_L(\tilde{x}_t, t)$$

and we may rewrite equation (28) as

$$\sigma_{BS}(K,T)^2 \approx \frac{1}{T}\int_0^T v_L(\tilde{x}_t)\,dt \tag{34}$$

In words, equation (34) says that the Black-Scholes implied variance of an option with strike K is given approximately by the integral from valuation date ($t = 0$) to the expiration date ($t = T$) of the local variances along the path $\tilde{x}_t$ that maximizes the Brownian Bridge density $q\,(x_t, t; x_T, T)$.

Of course in practice, it's not easy to compute the path $\tilde{x}_t$. Nevertheless, we now have a very simple and intuitive picture for the meaning of Black-Scholes implied variance of a European option with a given strike and expiration: it is approximately the integral from today to expiration of local variances along the most probable path for the stock price conditional on the stock price at expiration being the strike price of the option.

5 The Structure of Implied Volatility in the Heston Model

5.1 Local Volatility in the Heston Model

From Section 3.1 with $x_t \equiv Log\,(S(t)/K)$ and $\mu = 0$, we have

$$\begin{aligned} dx_t &= -\frac{v_t}{2}dt + \sqrt{v_t}\,dZ_t \\ dv_t &= -\lambda(v_t - \bar{v})dt + \rho\eta\sqrt{v_t}\,dZ_t + \sqrt{1-\rho^2}\eta\sqrt{v_t}\,dW_t \end{aligned} \tag{35}$$

where dW_t and dZ_t are orthogonal. Eliminating $\sqrt{v_t}dZ_t$, we get

$$dv_t = -\lambda(v_t - \bar{v})dt + \rho\eta\left(dx_t + \frac{1}{2}v_t dt\right) + \sqrt{1-\rho^2}\eta\sqrt{v_t}dW_t \tag{36}$$

Our strategy will be to compute local variances in the Heston model and then integrate local variance from valuation date to expiration date to get the BS implied variance following the results of Section 4.

First, consider the unconditional expectation $\hat{v}_s$ of the instantaneous variance at time s. Solving equation (36) gives

$$\hat{v}_s = (v_0 - \bar{v})e^{-\lambda s} + \bar{v}$$

Then define the expected total variance to time t through the relation

$$\hat{w}_t \equiv \int_0^t \hat{v}_s ds = (v_0 - \bar{v}) \left\{ \frac{1 - e^{-\lambda t}}{\lambda} \right\} + \bar{v}\, t$$

Finally, let $u_t \equiv \mathbb{E}[v_t \,| x_T]$ be the expectation of the instantaneous variance at time t conditional on the final value x_T of x.

Ansatz

(By "ansatz", I mean some working assumption which I haven't been able to justify and may not even be true). Without loss of generality, assume $x_0 = 0$. Then,

$$\mathbb{E}[x_s \,| x_T] = x_T \frac{\hat{w}_s}{\hat{w}_T}$$

where $\hat{w}_t \equiv \int_0^t ds\, \hat{v}_s$ is the expected total variance to time t . To see that this ansatz is plausible, note that

$$\mathbb{E}\,(x_s) = \mathbb{E}(x_T) \frac{\hat{w}_s}{\hat{w}_T} = -\frac{\hat{w}_T}{2} \frac{\hat{w}_s}{\hat{w}_T} = -\frac{\hat{w}_s}{2}$$

In fact, if the process for x_t were a conventional Brownian Bridge process, the result would be true but in this case, the result is only approximately true. If you manage to derive the correct result, please let me know.

Assuming the ansatz to be correct, we may take the conditional expectation of (36) to get:

$$du_t = -\lambda(u_t - \bar{v})dt + \frac{\rho\eta}{2} u_t dt + \rho\eta \frac{x_T}{\hat{w}_T} d\hat{w}_t + \sqrt{1-\rho^2}\eta\sqrt{v_t}\,\mathbb{E}[dW_t \,| x_T] \quad (37)$$

If the dependence of dW_t on x_T is weak or if $\sqrt{1-\rho^2}$ is very small, we may drop the last term to get

$$du_t \approx -\lambda'(u_t - \bar{v}')dt + \rho\eta \frac{x_T}{\hat{w}_T} \hat{v}_t dt$$

with $\lambda' = \lambda - \rho\eta/2$, $\bar{v}' = \bar{v}\lambda/\lambda'$. The solution to this equation is

$$u_T \approx \hat{v}'_T + \rho\eta \frac{x_T}{\hat{w}_T} \int_0^T \hat{v}_s\, e^{-\lambda'(T-s)} ds \quad (38)$$

with $\hat{v}'_s \equiv (v - \bar{v}')\, e^{-\lambda' s} + \bar{v}'$.

From Section 2.5, we know that the local variance $\sigma^2(K, T, S_0) = \mathbb{E}\left[v_T \left| S_T = K \right. \right]$. Then, equation (38) gives us an approximate but surprisingly accurate formula for local variance within the Heston model (an extremely accurate approximation when $\rho = \pm 1$) . We see that in the Heston model, local variance is approximately linear in $x = \log\left(\frac{F}{K}\right)$.

In summary, we have made two approximations: the Ansatz and dropping the last term in equation (37). For reasonable parameters, equation (38) gives good intuition for the functional form of local variance and when $\rho = \pm 1$, it is almost exact. Appendix A has a proof [2] that equation (38) is in fact exact to first order in η whether or not the ansatz holds or $\sqrt{1-\rho^2}$ is small.

5.2 Implied Volatility in the Heston Model

Now, to get implied variance in the Heston model, following the results of Section 4, we need to integrate the Heston local variance along the most probable stock price path joining the initial stock price to the strike price at expiration (the one which maximizes the Brownian Bridge probability density).

In the notation of Section 4, the Black-Scholes implied variance is given by

$$\sigma_{BS}(K,T)^2 \approx \frac{1}{T}\int_0^T \sigma^2_{\tilde{x}_t,t}\, dt = \frac{1}{T}\int_0^T u_t(\tilde{x}_t)\, dt \tag{39}$$

where $\{\tilde{x}_t\}$ is the most probable path (as defined above).

Recall that the Brownian Bridge density $q\,(x_t, t; x_T, T)$ is roughly symmetric and peaked around $\tilde{x}_t$, so $\mathbb{E}\left[x_t - \tilde{x}_t \left| x_T \right. \right] \approx 0$. Applying the Ansatz once again, we obtain

$$\tilde{x}_t = \mathbb{E}\left[\tilde{x}_t \left| x_T \right. \right] = \mathbb{E}\left[\tilde{x}_t - x_t \left| x_T \right. \right] + \mathbb{E}\left[x_t \left| x_T \right. \right] \approx \frac{\hat{w}_t}{\hat{w}_T} x_T$$

We substitute this expression back into equations (38) and (39) to get

$$\begin{aligned} \sigma_{BS}(K,T)^2 &\approx \frac{1}{T}\int_0^T u_t(\tilde{x}_t)\, dt \\ &\approx \frac{1}{T}\int_0^T \hat{v}'_t dt + \rho\eta \frac{x_T}{\hat{w}_T}\frac{1}{T}\int_0^T dt \int_0^t \hat{v}_s\, e^{-\lambda'(t-s)} ds \end{aligned} \tag{40}$$

[2] Thanks to Peter Friz for pointing this out and proving it.

The BS Implied Volatility Term Structure in the Heston Model

The at-the-money term structure of BS implied variance in the Heston model is obtained by setting $x_T = 0$ in equation (40). Performing the integration explicitly gives

$$\begin{aligned}\sigma_{BS}(K,T)^2\big|_{K=F_T} \approx \frac{1}{T}\int_0^T \hat{v}'_t dt &= \frac{1}{T}\int_0^T \left[(v-\bar{v}')\,e^{-\lambda' t} + \bar{v}'\right] dt \\ &= (v-\bar{v}')\,\frac{1-e^{-\lambda' T}}{\lambda' T} + \bar{v}'\end{aligned}$$

We see that in the Heston model, the at-the-money Black-Scholes implied variance $\sigma_{BS}(K,T)^2|_{K=F_T} \to v$ (the instantaneous variance) as the time to expiration $T \to 0$ and as $T \to \infty$, the at-the-money Black-Scholes implied variance reverts to $\bar{v}'$.

The BS Implied Volatility Skew in the Heston Model

It is possible (but not very illuminating) to integrate the second term of equation (40) explicitly. Even without doing that, we can see that the implied variance skew in the Heston model is approximately linear in the correlation ρ and the volatility of volatility η.

In the special case where $v_0 = \bar{v}$, the implied variance skew has a particularly simple form. Then $\hat{v}_s = \bar{v}$ and $\hat{w}_t = \bar{v}\,t$. The most probable path $\tilde{x}_t \approx \frac{t}{T} x_T$ is exactly a straight line in log-space between the initial stock price on valuation date and the strike price at expiration. Performing the integrations in equation (40) explicitly, we get

$$\begin{aligned}\sigma_{BS}(K,T)^2 &\approx \frac{\hat{w}'_T}{T} + \rho\eta\frac{x_T}{T^2}\int_0^T dt\,\frac{1}{T}\int_0^t e^{-\lambda'(t-s)}ds \\ &= \frac{\hat{w}'_T}{T} + \rho\eta\,\frac{x_T}{\lambda' T}\left\{1 - \frac{\left(1-e^{-\lambda' T}\right)}{\lambda' T}\right\}\end{aligned} \tag{41}$$

From equation (41), we see that the implied variance skew $\frac{\partial}{\partial x_t}\sigma_{BS}(K,T)^2$ is *independent* of the level of instantaneous variance v or long-term mean variance $\bar{v}$. In fact, this remains approximately true even when $v \neq \bar{v}$. It follows that we now have a fast way of calibrating the Heston model to observed implied volatility skews. Just two expirations would in principle allow us to determine λ' and the product $\rho\eta$. We can then fit the term structure of volatility to determine the long term mean variance $\bar{v}$ and the instantaneous variance v_0. The curvature of the skew (not discussed here) would allow us to determine ρ and η separately.

We note that as we increase either the correlation ρ or the volatility of volatility η, the skew increases.

Also, the very short-dated skew is independent of λ and T:

$$\frac{\partial}{\partial x_t}\sigma_{BS}(K,T)^2 = \rho\eta\,\frac{1}{\lambda' T}\left\{1-\frac{\left(1-e^{-\lambda' T}\right)}{\lambda' T}\right\} \to \frac{\rho\eta}{2} \quad \text{as} \quad T\to 0$$

and the long-dated skew is inversely proportional to T:

$$\frac{\partial}{\partial x_t}\sigma_{BS}(K,T)^2 = \rho\eta\,\frac{1}{\lambda' T}\left\{1-\frac{\left(1-e^{-\lambda' T}\right)}{\lambda' T}\right\} \sim \frac{\rho\eta}{\lambda' T} \quad \text{as} \quad T\to \infty$$

Finally, increasing η causes the curvature of the implied volatility skew (related to the kurtosis of the risk-neutral density) to increase but we haven't shown that here.

6 The SPX Implied Volatility Surface

Up to this point, we have concentrated on understanding the shape of the implied volatility surface implied by a stochastic volatility model – in particular the Heston model. However, we still have no idea whether implied volatilities produced by the Heston model look like implied volatilities in the market. For reference, Appendix B has graphs of the SPX implied volatility surface for 3 days prior to the Sep-02 expiration.

After performing a nonlinear fit to observed variance as a function of x_t for each expiration t, we get the at-the-money forward variance levels and skews listed in Table 1. (Recall that by at-the-money skew, I mean $\frac{\partial}{\partial x_t}\sigma_{BS}(K,T)^2$).

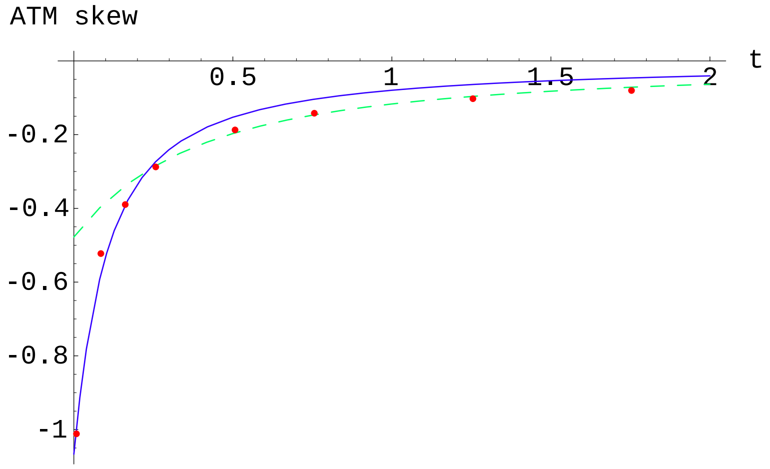
Skew is plotted as a function of time in Figure 2. Just looking at the pattern of the points, we would suspect that a simple functional form should be able to fit. However, the solid and dashed lines show the results of fitting the approximate formula

$$\rho\eta\frac{1}{\lambda' T}\left\{1-\frac{\left(1-e^{-\lambda' T}\right)}{\lambda' T}\right\}$$

Table 1: At-the-money SPX variance levels and skews on September 17, 2002.

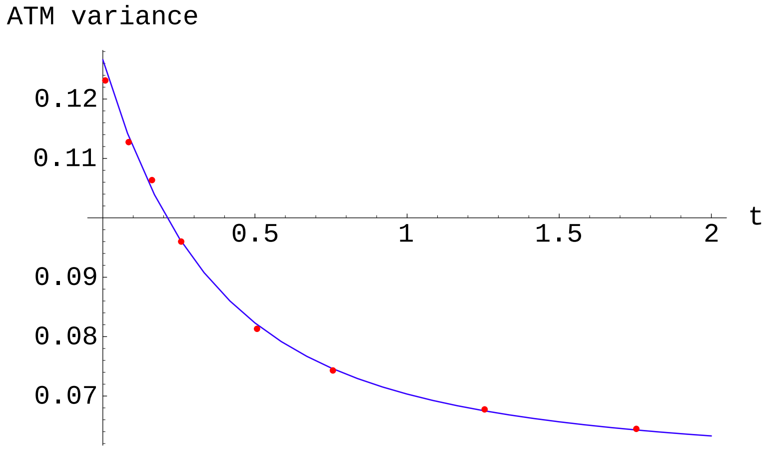
Expiration	**Time (years)**	**ATM Variance**	**ATM Skew**
Sep-02	0.0082	0.1231	-1.0115
Oct-02	0.0849	0.1127	-0.5226
Nov-02	0.1616	0.1064	-0.3897
Dec-02	0.2575	0.0960	-0.2876
Mar-03	0.5068	0.0813	-0.1872
Jun-03	0.7562	0.0743	-0.1419
Dec-03	1.2548	0.0677	-0.1025
Jun-04	1.7534	0.0645	-0.0803

Figure 2: Graph of SPX ATM skew vs. time. The dashed fit excludes the first 3 data points.



to the observed skews. The solid line takes all points into account; the dashed line drops the first three expirations from the fit. We can see that the fitting function is too stiff to fit the observed pattern of variance skews; there is no choice of λ' that will allow us to fit the skew observations. The fact that the observed variance skew increases significantly faster as $T \to 0$ than the skew implied by a stochastic volatility model may indicate that jumps need to be included in a complete model as in Matytsin (1999) for example. We will explore this further in the next lecture.

Figure 3: Graph of SPX ATM variance vs. time.



In Figure 3, we see that on this particular date, our simple formula fits the data beautifully. It should be emphasized that this is not always the case; in general, the term structure of volatility can be quite intricate at the short-end.

$$\sigma_{BS}(K,T)^2\big|_{K=F_T} \approx (v - \bar{v}')\,\frac{1}{\lambda' T}\left\{1 - \frac{\left(1 - e^{-\lambda' T}\right)}{\lambda' T}\right\} + \bar{v}'$$

So, sometimes it's possible to fit the term structure of at-the-money volatility with a stochastic volatility model, but it's never possible to fit the term structure of the volatility skew for short expirations. Now we understand one reason why practitioners prefer local volatility models – a stochastic volatility model with time-homogeneous parameters cannot fit market prices! Perhaps an extended stochastic volatility model with correlated jumps in stock price and volatility (Matytsin (1999)) might fit better? But how would traders choose their input parameters? How would the SPX index book trader choose his volatility of volatility parameter – or worse, the correlation between jumps in stock price and jumps in volatility?

References

Avellaneda, Marco, C. Friedman, R. Holmes, and D. Samperi, 1997, Calibrating volatility surfaces via relative-entropy minimization, *Applied Mathematical Finance* 4, 37–64.

Blacher, Guillaume, 1998, Local volatility, RISK Conference Presentation.

Brigo, Damiano, and Fabio Mercurio, 2002, Lognormal-mixture dynamics and calibration to market volatility smiles, *International Journal of Theoretical and Applied Finance* 5, 427–446.

Derman, Emanuel, Iraj Kani, and Neil Chriss, 1996, Implied trinomial trees of the volatility smile, *Journal of Derivatives* 3, 7–22.

Lee, Roger W., 2002, Implied volatility: Statics, dynamics, and probabilistic interpretation, Discussion paper Department of Mathematics, Stanford University.

Matytsin, Andrew, 1999, Modelling volatility and volatility derivatives, Columbia Practitioners Conference on the Mathematics of Finance.

Mikhailov, Sergei, and Ulrich Nögel, July 2003, Heston's stochastic volatility model, calibration and some extensions, *Wilmott Magazine* pp. 74–79.

Rubinstein, Mark, 1998, Edgeworth binomial trees, *Journal of Derivatives* 5, 20–27.

Shimko, D., 1993, Bounds on probability, *Risk* 6, 33–37.

A Proof that equation (38) is correct to first order in η

The following proof is due to Peter Friz.

We begin by expanding equation (38) to first order in η to get

$$u_T \approx \hat{v}_T + \rho\eta \left\{ \frac{\bar{v}}{2\lambda}(1 - e^{-\lambda T}) - \frac{T}{2} e^{-\lambda T}(v_0 - \bar{v}) + \frac{x_T}{\hat{w}_T} \int_0^T e^{-\lambda(T-s)} \hat{v}_s ds \right\} \tag{A-1}$$

Next, recall that local variance is given by $E[v_T|x_T]$. Writing x_t and v_t explicitly as

$$x_t = x_t^{(0)} + \eta x_t^{(1)} + o(\eta).$$

and

$$v_t = v_t^{(0)} + \eta v_t^{(1)} + o(\eta)$$

(note $v_t^{(0)} = \hat{v}_t$) and expanding to first order in η gives

$$\begin{aligned} E[v_T|x_T] &= E[v_T^{(0)} + \eta v_T^{(1)} | x_T^{(0)} + \eta x_T^{(1)}] + o(\eta) \\ &= \hat{v}_T + \eta E[v_T^{(1)} | x_T^{(0)}] + o(\eta) \end{aligned} \tag{A-2}$$

since $\hat{v}_T = v_T^{(0)}$ is deterministic. So, to compute the local variance to first order in η, we need only compute $E[v_T^{(1)} | x_T^{(0)}]$.

By formally differentiating the Heston SDEs (35) w.r.t. η at $\eta = 0$ and with $d\tilde{W}_t$, we obtain

$$\begin{aligned} dx_t^{(0)} &= -\frac{1}{2} v_t^{(0)} dt + \sqrt{v_t^{(0)}} \left(\rho\, dW_t + \sqrt{1-\rho^2} dZ_t \right) \\ dv_t^{(1)} &= -\lambda v_t^{(1)} dt + \sqrt{v_t^{(0)}} dW_t \end{aligned}$$

with initial conditions $x_0^{(0)} = v_0^{(1)} = 0$.

The solutions to these SDEs are

$$x_T^{(0)} = -\frac{1}{2} \hat{w}_T + \rho \int_0^T \sqrt{\hat{v}_s} dW_s + \sqrt{1-\rho^2} \int_0^T \sqrt{\hat{v}_s} dZ_s.$$

and

$$v_T^{(1)} = \int_0^T e^{-\lambda(T-s)} \sqrt{\hat{v}_s}\, dW_s.$$

We note that both $x_T^{(0)}$ and $v_T^{(1)}$ are Gaussian random variables so to compute $E[v_T^{(1)}|x_T^{(0)}]$, we need to know how to compute the expectation of a Gaussian random variable conditional on another Gaussian random variable. For this we have the following lemma:

Lemma 1. (Normal regression) Let $\tilde{X}$,$\tilde{Y}$ be zero mean Gaussian random variables. Then

$$E[\tilde{X}|\tilde{Y}] = \tilde{Y}\frac{\mathrm{Cov}[\tilde{X},\tilde{Y}]}{V[\tilde{Y}]}$$

For non-zero mean Gaussians X, Y this extends to

$$E[X|Y] = Y\frac{\mathrm{Cov}[X,Y]}{V[Y]} + \frac{E[X]V[Y] - E[Y]\mathrm{Cov}[X,Y]}{V[Y]}$$

Example (Brownian Bridge): $E[B_t|B_T] = B_T\frac{t}{T}$.

Applying this lemma, we obtain

$$\begin{aligned} E[v_T^{(1)}|x_T^{(0)} = x] &= (x - E[x_T^{(0)}])\,\frac{\mathrm{Cov}[v_T^{(1)}, x_T^{(0)}]}{V[x_T^{(0)}]} \\ &= \left(x + \frac{1}{2}w_T\right)\,\frac{\rho\int_0^T e^{-\lambda(t-s)}\hat{v}_s\,ds}{\hat{w}_T} \end{aligned}$$

Substituting this result into equation (A-2) gives

$$\begin{aligned} E[v_T|x_T] &= \hat{v}_T + \rho\eta\left(x + \frac{1}{2}\hat{w}_T\right)\,\frac{\int_0^T e^{-\lambda(T-s)}\hat{v}_s\,ds}{\hat{w}_T} + o(\eta) \\ &= \hat{v}_T + \frac{\rho\eta}{2}\int_0^T e^{-\lambda(T-s)}\hat{v}_s\,ds + \rho\eta\frac{x}{\hat{w}_T}\int_0^T e^{-\lambda(T-s)}\hat{v}_s\,ds + o(\eta) \end{aligned} \tag{A-3}$$

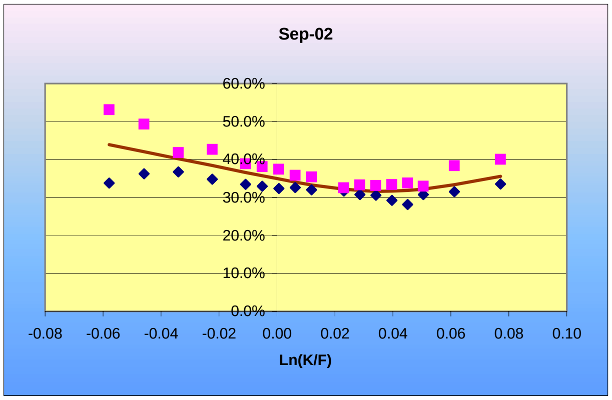
It's easy to check that

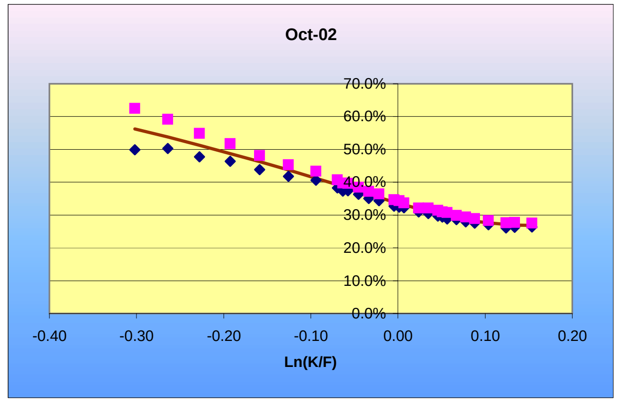
$$\int_0^T e^{-\lambda(T-s)}\hat{v}_s\,ds = \frac{\bar{v}}{2\lambda}(1 - e^{-\lambda T}) - \frac{T}{2}e^{-\lambda T}(v_0 - \bar{v})$$

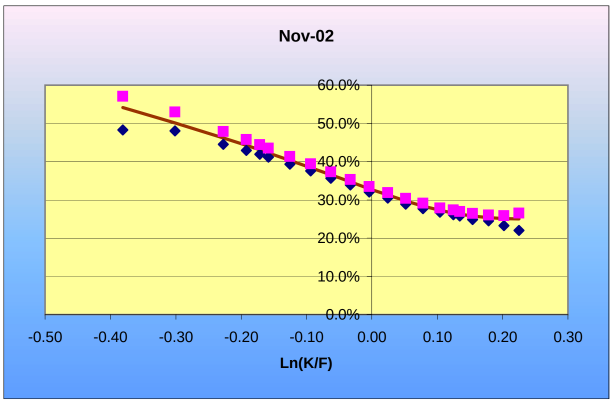
Then equations (A-1) and (A-3) are identical. We conclude that equation (38) is true up to $o(\eta)$ *whether or not* the ansatz holds or $\sqrt{1-\rho^2}$ is small.

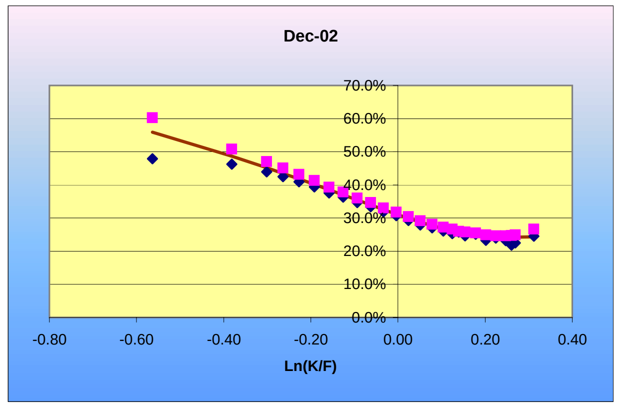
B SPX Volatility Surfaces for the 3 Days prior to the Sep-02 Expiration

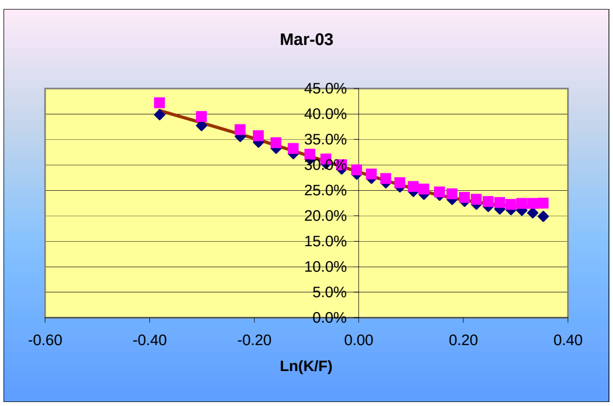
SPX implied volatility as a function of x=Ln(K/F) as of 17-Sep-2002

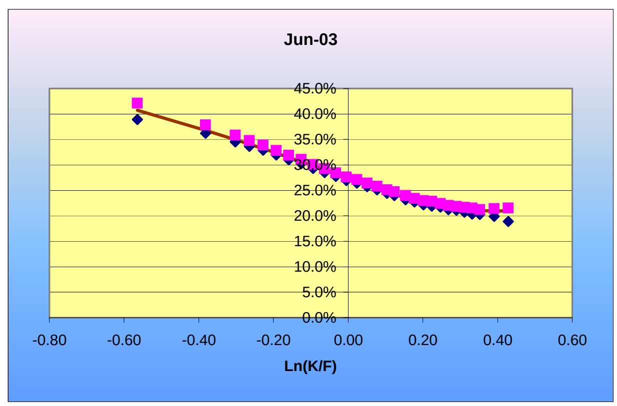


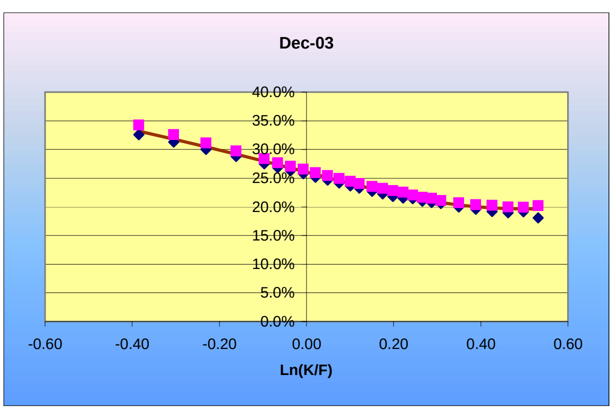


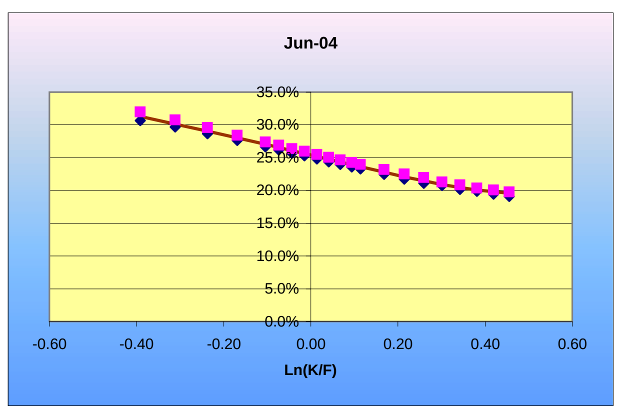




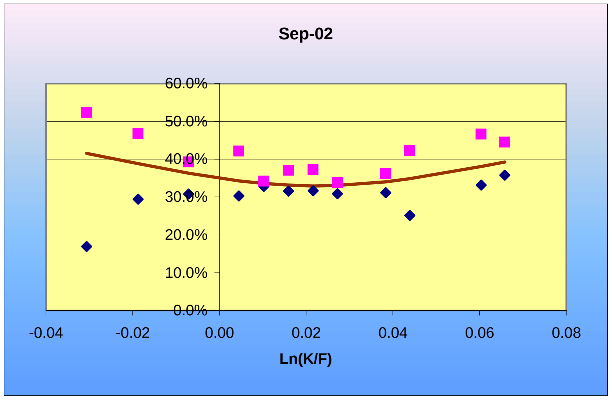


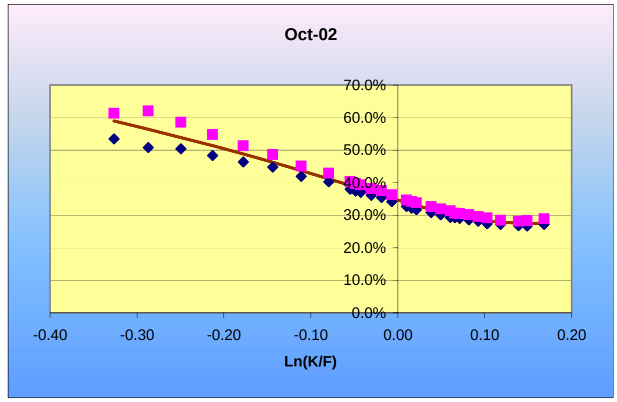


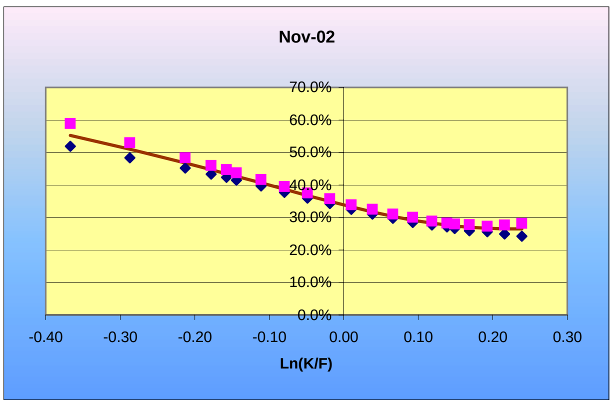


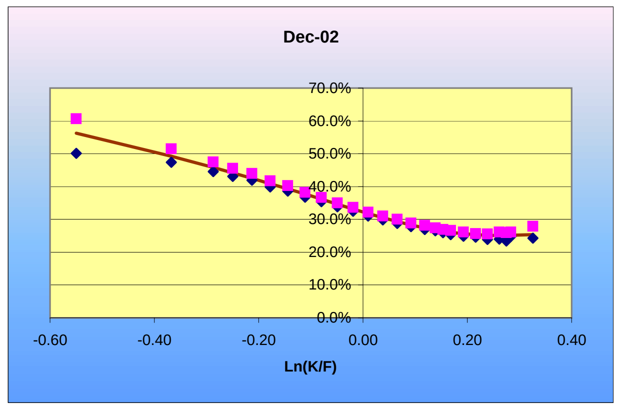


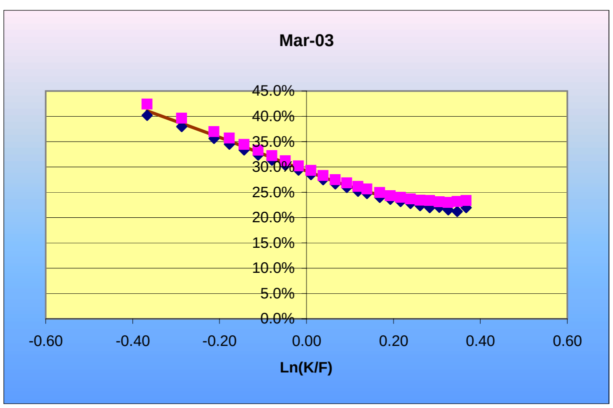
SPX implied volatility as a function of x=Ln(K/F) as of 18-Sep-2002

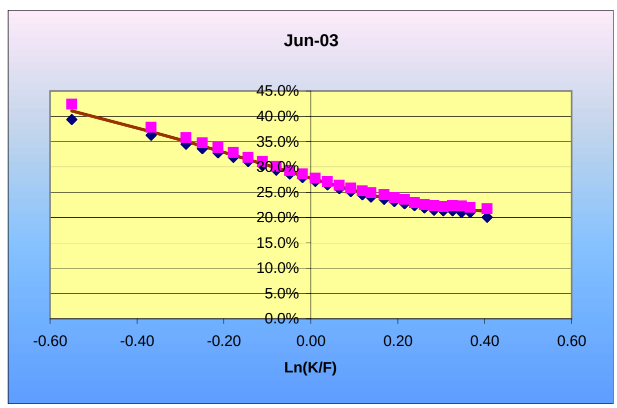


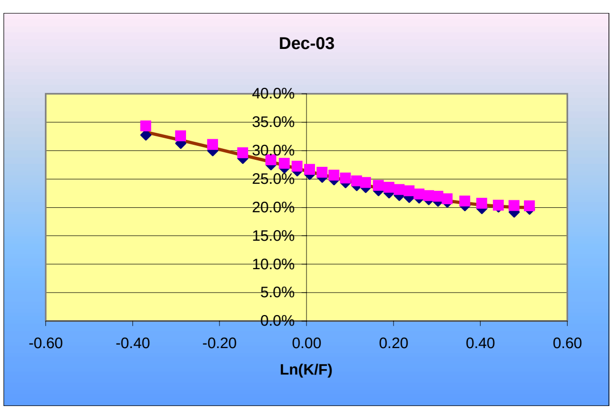


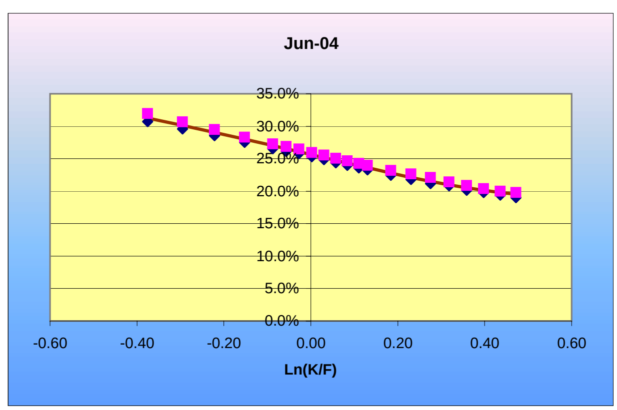












SPX implied volatility as a function of x=Ln(K/F) as of 19-Sep-2002