Beta Stochastic Volatility Model

Piotr Karasinski and Artur Sepp *

September 22, 2012

Published in:
Risk Magazine, October 2012, 66-71

Abstract

We introduce the beta stochastic volatility model and discuss empirical features of this model and its calibration. This model is appealing because, first, its parameters can be easily understood and calibrated and, second, it produces steeper forward skews, compared to traditional stochastic volatility models.

1 Introduction

The traditional approach to stochastic volatility (SV) modeling begins with specification of a SV process, most typically on the sole grounds of its analytical tractability (for example, Heston (1993)). Then, after a closed-form solution for vanilla options has been derived and implemented, parameters of the SV process are calibrated to the implied volatility surface of vanilla options using complicated non-linear optimization methods. The drawback of this approach for business applications is that, on one hand, the calibrated parameters are unstable from day to day and their change cannot be explained using changes in market observables for the implied volatility surface (at-the-money (ATM) volatility, skew, convexity). On the other hand, key variables, such as forward skews and volatility of volatility, are outputs in these models and cannot be tweaked to reproduce market observables or proprietary views.

The problem in modeling of forward skews was relaxed by the introduction of local stochastic volatility models (Lipton (2002)) which combine the local volatility dynamics, to fit the vanilla surface by construction, and the SV dynamics, to model forward skews. The advantage of local stochastic volatility (LSV) models is that, since the vanilla surface is matched by construction for any admissible set of model parameters for the SV dynamics, the modeler has the freedom to specify parameters of the SV process to match either historical data or market data for exotic options. The typical driver for the SV part is chosen to be either the square root process, as in the Heston model, or exponential Ornstein-Uhlenbeck process.

*Piotr Karasinski is a senior adviser at the European Bank for Reconstruction and Development in London. Artur Sepp is a vice-president of the global quantitative group at Bank of America Merrill Lynch in London. They are grateful to their colleagues and two anonymous referees for helpful and instructive comments. Artur is particularly grateful to Hassan El Hady for bringing to his attention some early work on this model by Piotr. Email: piotr.g.karasinski@gmail.com, artur.sepp@baml.com

Electronic copy available at: http://ssrn.com/abstract=2150614

While LSV models became popular in foreign exchange markets, in equity markets they are less popular. One disadvantage of LSV models is their tendency to under-price volatility products (such as cliquets) because they imply high convexity for forward smile and, as a result, overpricing caps and underpricing floors of forward-starting products. This can be relaxed by introducing jumps in the model dynamics (Sepp (2011)), although at the expense of extra model parameters - one for the jump intensity and two for jump sizes in the price and volatility. Another disadvantage is the difficulty in interpreting the SV parameters (equity-volatility correlation, volatility-of-volatility and mean-reversion rate, not counting for parameters of jump process) and their impact on forward skews.

In this article, we introduce the beta stochastic volatility model to overcome above mentioned disadvantages. The key parameter of the model, β, can be naturally interpreted as the rate of change in the short-term ATM volatility given change in the spot price and easily implied from historical and current market data. The remaining two parameters - the idiosyncratic volatility of volatility and the mean-reversion rate have lesser impact on forward skews and also can be estimated from historical data without the need to apply time-expensive and obscure non-linear fits. The beta stochastic volatility for the logarithm of the volatility was first implemented by Bardhan-Karasinski (1993). A similar model, but in the discrete time setting, was developed by Langnau (2004). Here we describe the model dynamics incorporating the local volatility with calibration that can be implemented using conventional partial-differential equations (PDE) methods. In the limiting case of the deterministic local volatility, we derive an accurate approximation for prices of vanilla options in this model. Finally, we provide illustrations of model calibration and its implications.

2 The dynamics

We consider two specifications of model dynamics for the underlying price $S(t)$ and its instantaneous stochastic volatility. The first one is an econometric version applied for model analysis using time series, the second one is pricing version applied for risk-neutral valuation purposes in the context of local stochastic volatility models.

The econometric model is expressed on terms of stochastic volatility process $V(t)$:

$$\begin{aligned}
&\frac{dS(t)}{S(t)} = \widehat{\mu}(t)dt + V(t)dW^{(0)},\ S(0) = S \\
&dV(t) = \widehat{\kappa}(\widehat{\theta} - V(t))dt + \widehat{\beta}V(t)dW^{(0)} + \widehat{\varepsilon}dW^{(1)},\ V(0) = V_0
\end{aligned} \tag{2.1}$$

where:
$\widehat{\beta}$ ($\widehat{\beta} < 0$) is the rate of change in the volatility corresponding to change in the spot price;
$\widehat{\varepsilon}$ is idiosyncratic volatility of volatility ;
$\widehat{\kappa}$ is the mean-reversion rate;

Electronic copy available at: http://ssrn.com/abstract=2150614

$\widehat{\theta}$ is the mean level of volatility;
$W^{(0)}(t)$ and $W^{(1)}(t)$ are two Brownians with $dW^{(0)}dW^{(1)} = 0$;
$\widehat{\mu}(t)$ is the drift.

The pricing model is expressed in terms of the normalized volatility factor $Y(t)$:

$$\begin{aligned} \frac{dS(t)}{S(t)} &= \mu(t)dt + \sigma(1 + Y(t))dW^{(0)}, \ S(0) = S \\ dY(t) &= -\kappa Y(t)dt + \beta\sigma(1 + Y(t))dW^{(0)} + \varepsilon dW^{(1)}, \ Y(0) = 0 \end{aligned} \tag{2.2}$$

where:
σ is the overall level of volatility - we assume that σ is set to either constant volatility σ_{CV} or deterministic volatility $\sigma_{DV}(t)$, or local stochastic volatility $\sigma_{LSV}(t, S)$: $\sigma = \{\sigma_{CV}, \sigma_{DV}(t), \sigma_{LSV}(t, S)\}$;
$W^{(0)}(t)$ and $W^{(1)}(t)$ are two Brownians with $dW^{(0)}dW^{(1)} = 0$;
$\mu(t)$ is the risk-neutral drift;

For constant volatility σ_{CV}, the following relationship holds between parameters of the econometric and pricing models:

$$Y(t) = \frac{V(t)}{\sigma_{CV}} - 1 \ , \ \beta = \frac{\widehat{\beta}}{\sigma_{CV}} \ , \ \varepsilon = \frac{\widehat{\varepsilon}}{\sigma_{CV}} \ , \ \kappa = \widehat{\kappa} \tag{2.3}$$

We note that the volatilities in the dynamics (2.2), for constant volatility σ_{CV}, are linear in $Y(t)$ so that the unique global solution exists for this pair of SDE-s (this might not be the case for more conventional exponential Ornstein-Uhlenbeck model), which is important for stability of Monte Carlo (MC) and PDE methods.

2.1 Intuition

To get intuition into the dynamics (2.1) of the econometric model, we assume no mean-reversion and idiosyncratic volatility of volatility, taking $\widehat{\kappa} = \widehat{\varepsilon} = 0$ and $\widehat{\mu}(t) = 0$. We assume that the dynamics of the ATM implied volatility for an infinitesimal maturity is specified by $\sigma_{ATM}(t) = V(t)$, so that using (2.1) we obtain:

$$d\sigma_{ATM}(t) = \widehat{\beta}\frac{dS(t)}{S(t)} \tag{2.4}$$

As a result, the dynamics (2.4) can be interpreted as a regression model with $\sigma_{ATM}(t)$ increasing (as $\widehat{\beta} < 0$) if the spot price decreases and vice-versa, producing the leverage effect observed in the market.

The coefficient $\widehat{\beta}$ can be estimated using the time series $\sigma_{ATM}(t)$ and $S(t)$ by means of standard regression methods as we describe below. Empirically, there is strong correlation between changes in $\sigma_{ATM}(t)$ and $S(t)$. In Figure 1 (top left), we show the scatter plot daily changes in 1m ATM volatility corresponding to daily returns in the S&P500 index using market data from October 2009 to October 2011, and the corresponding regression model. We see that the regression model explains about 80% of daily variations in $\sigma_{ATM}(t)$ so that the model has a strong explanatory power. The rest of the variation is modelled by the idiosyncratic volatility of volatility $\widehat{\varepsilon}$. Empirically, the residual is independent from $S(t)$ so that we can safely assume that the correlation between two Brownians in model dynamics (2.2) is zero to reduce the number of model parameters.

2.2 Comparison to the SABR and other stochastic volatility models

Now we consider the dynamics of the pricing model (2.2) with no mean-reversion and idiosyncratic volatility of volatility, $\kappa = \varepsilon = 0$. Using notations as in that SABR model by Hagan *et al* (2002) with $\beta = 1$, where β in the SABR model is the CEV beta, we obtain the following specification of SABR model parameters in terms of parameters of the beta stochastic volatility model:

$$\hat{a} = \sigma_{CV}(1 + Y(t)) \ , \ \rho = -1 \ , \ \nu = -\beta\sigma_{CV} \ , \ \alpha = \sigma_{CV} \tag{2.5}$$

In particular, using formula (3.1a) for a short maturity and small log-moneynes k, $k = \ln(K/S_0)$, we obtain the following relationship for the Black-Scholes-Merton implied volatility $\sigma_{IMP}(k)$:

$$\sigma_{IMP}(k) = \sigma_{CV} + \frac{1}{2}\sigma_{CV}\beta k - \frac{1}{12}\sigma_{CV}\left(\beta k\right)^2 \tag{2.6}$$

The main similarity between the beta SV model and the SABR model is that both assume the log-normal process for the instantaneous volatility (in the restricted case (2.5)). In general, the volatility process in the dynamics (2.2) is a mixture of log-normal and normal components.

The difference between the two models is that, in the SABR model (log-normal case), the covariance between spot and volatility moves ($\nu\hat{a}\rho dW^0$) and the variance of idiosyncratic modes in the volatility ($\nu\hat{a}\sqrt{1-\rho^2}dW^1$) are modeled through the same set of parameters. In the beta SV model, these are modeled through independent parameters so that we have more freedom in modeling the skew and curvature of the implied volatility.

Importantly, the key difference between the beta SV model and the SABR model is that the SABR model is mostly used for the parametrization of the implied volatility surface while, to use a stochastic volatility model for pricing path-dependent options, we require a mean-reversion. The intuition behind the mean-reversion is that when the ATM volatility is small then it will tend to rise while when it is too large is will tend to decline. Thus, over long time horizon (say 2y or 3y), the realized volatility of volatility is expected to stay in a range and not increase with maturity time, as implied by models with no mean-reversion. The concept of mean-reversion is related with the existence of the steady-state distribution and the ergodicity of the process (the long-term distribution does not depend on maturity time). The steady-state distribution of the volatility does not exist in the SABR model, however it does exist in the beta SV model - see equation (5.7).

2.3 Interpretation of parameter β

Now following the same assumptions as in the previous section, we use formula (2.6) as an alternative way to calibrate the parameter β from the implied market volatility $\sigma_{IMP}(k)$ and skew:

$$\beta = \frac{\text{Skew}}{\sigma_{IMP}(0)}\ ,\ \text{Skew} = \frac{\sigma_{IMP}(k_+) - \sigma_{IMP}(k_-)}{\frac{1}{2}(k_+ - k_-)} \tag{2.7}$$

where k_+ and k_-, $k_+ \approx -k_-$, are the log-strikes for OTM call and put, respectively (typically, 10% strikes are used). Thus, we obtain that the model implies the following approximate relationship:

$$\text{Skew}(t) = \beta\sigma_{ATM}(t) \tag{2.8}$$

This relationship is observed empirically. In Figure 1 (top right) we show the scatter plot of 1m 110%-90% skew (unscaled) against 1m ATM volatility. The relationship is relatively strong, in particular, high ATM volatilities are associated with steep skews.

We note that in the Heston SV model, the short-term skew is approximately given by $\rho\epsilon/(2\sqrt{V(0)})$, where ρ is the correlation between the variance V and log-returns and ϵ is the volatility of variance (Bergomi (2004)). Thus, the Heston model implies the inverse relationship between short-term ATM volatility and skew. As a result, by model calibration in the presence of steep skews and high ATM volatilities, the Heston model implies very high volatility of variance ϵ (see Figure 2 in Bergomi (2004)), as $\rho \approx -1$ and the only way for the model to fit the skew is by increasing ϵ. In opposite, in the beta SV model, idiosyncratic volatility of volatility ε is stable over different market conditions.

We finally note that (2.4) and (2.8) imply the following approximate dynamics for the short-term skew in the econometric model:

$$d\text{Skew}(t) = \widehat{\beta}^2 \frac{dS(t)}{S(t)} \tag{2.9}$$

Thus, the model produces explicit dependence between changes in the skew given changes in the spot price, unlike conventional SV models. The dependence between changes in the skew and spot returns is observed empirically. In Figure 1 (bottom), we show the scatter plot daily changes in 1m 110%-90% skew corresponding to daily changes in the S&P500 index, and the corresponding regression model. Although it is less strong as that for the ATM volatility, the corresponding regression model explains 33% of daily variations in the skew. The slope is positive as negative changes in the S&P500 increase hedging demand for OTM puts thus increasing the skew.

3 Estimation of model parameters using time series

First, we discuss how to estimate model parameters $\widehat{\beta}$, $\widehat{\kappa}$, and $\widehat{\varepsilon}$ and examine the time series of these parameters. For this purpose, we use the time series of daily returns $\{R(t_i)\}$, $R(t_i) = \frac{S(t_i)}{S(t_{i-1})} - 1$, on the S&P 500 index and the VIX, which we use as a proxy for 1m ATM implied volatilities, $\{\sigma_{ATM}(t_i)\}$. We assume that $V(t_i)$ is proxies by the ATM volatility $V(t_i) = \sigma_{ATM}(t_i) \equiv \mathrm{VIX}(t_i)$.

The continuous time dynamics (2.2) can be approximated in the discrete time using the following regression model:

$$V(t_i) - V(t_{i-1}) = \widehat{\beta} R(t_i) + \eta \left(\overline{\sigma}_{ATM} - V(t_{i-1})\right) + \vartheta(t_i) \tag{3.1}$$

where $\overline{\sigma}_{ATM}$ is the average of $\{\sigma_{ATM}(t_i)\}$, $\eta \equiv \widehat{\kappa} dt$ with $dt = 1/252$ and $\vartheta(t_i)$ are independent normally distributed residuals. The standard deviation of residuals $\{\vartheta(t_i)\}$ in regression (3.1) is used to estimate $\widehat{\varepsilon}$ (applying the annualization factor $\sqrt{252}$).

For illustration, we use the time series of the S&P 500 index and VIX from January 1990 to January 2012. For each year in the sample, we apply the regression model (3.1) for daily returns to estimate the model parameters using data corresponding to given year. Our results are reported in Table 1. Here, we use the following notations: "SP500" and "VIX" denote averages of the S&P 500 and VIX closing levels, respectively, during the given year; "R SP500" with "St SP500" and "R VIX" with "St VIX" stand for the average daily return with standard deviation, both annualized, on the S&P 500 index (arithmetic return) and the VIX (absolute return), respectively; "Beta" and "Kappa" are estimates for parameters $\widehat{\beta}$ and $\widehat{\kappa}$; "St Resid" stands for standard deviation of residuals, annualized; R^2 is the explanatory power of the regression; "Correl" denotes correlation between returns on the S&P 500 index and VIX; "Skew" is the implied skew computed by $10\%\widehat{\beta}\overline{\sigma}_{ATM}$. Finally, "State" stands for the market regime with three states: "Risk Off" and "Risk On" when average annual returns on the S&P 500 index are negative and positive respectively, "Range" when returns are within 3%.

We see that the VIX increases during "Risk Off" years and declines during "Risk On", while the skew typically declines right before "Risk Off" or during "Range" years, when the demand for protective OTM puts increases relative for ATM options. In contrast, the VIX increases during "Risk Off" years, when demand for both ATM puts and calls increases relative for OTM puts, as latter become prohibitively expensive given high levels of the ATM volatility.

We see that $\widehat{\kappa}$ increases during "Risk On" years as any shock in volatility dissipates fast driven by "buy the dip" mentality, while it is very low right after "Risk Off" or "Range" years implying that the volatility declines very slowly from elevated levels given the prevalent risk-aversion. At the same time, $\widehat{\beta}$ increases during "Risk On" years due to the same reasons we provided for behavior of the skew. We see that, naturally, the standard deviations increase during "Risk Off" years. Remarkably, the idiosyncratic volatility $\widehat{\varepsilon}$ is relatively stable ranging between $10\% - 20\%$, this behavior is to be contrasted with the volatility of volatility in the Heston model which is known to be highly unstable.

Finally, we note that the model R^2 tends to increase during "Risk On" and "Range" years while declines right after "Risk Off" years as declines in the volatility tend to lag increases in stock prices. Remarkably, we observe the increasing growth in R^2 and absolute value of the correlation during the last decade, in line with intensified correlations in cross-asset returns.

To conclude, we find that the model R^2 and idiosyncratic volatility $\widehat{\varepsilon}$ are the relative stable, while changes in $\widehat{\beta}$ and $\widehat{\kappa}$ can be explained by prevalent market conditions. For pricing applications, we can use a term structure of β to fit forward skews or reflect proprietary views on the market dynamics.

4 Pricing equation

The model dynamics (2.2) do not have a closed form solution for vanilla options even with a constant volatility. In appendix A, we obtain an approximate solution for call option values under a constant σ_{CV} or deterministic volatility $\sigma_{DV}(t)$. In Figure 4, we illustrate that the proposed formula is in good agreement with numerical PDE solution across different maturities and strikes, especially for near ATM strikes. Nevertheless, the calibration of local volatility can only be implemented by means of partial differential equation (PDE) methods. Here we derive the necessary equations.

Using dynamics (2.2), for the log-spot, $X(t) = \ln\left(\frac{S(t)}{S(0)}\right)$ we obtain:

$$\begin{aligned} dX(t) &= \mu(t)dt - \frac{1}{2}\sigma^2(1+Y(t))^2dt + \sigma(1+Y(t))dW^{(0)},\ X(0)=0 \\ dY(t) &= \beta\sigma(1+Y(t))dW^{(0)} - \kappa Y(t)dt + \varepsilon dW^{(1)},\ Y(0)=0 \end{aligned} \tag{4.1}$$

with

$$\begin{aligned} dY(t)dY(t) &= \left(\varepsilon^2 + \beta^2\sigma^2(1+Y(t))^2\right)dt \\ dX(t)dY(t) &= \beta\sigma^2(1+Y(t))^2dt \end{aligned}$$

Finally, the pricing equation for value function $U(t,T,X,Y)$ has the form:

$$\begin{aligned} &U_t + \frac{1}{2}\sigma^2(1+2Y+Y^2)\left[U_{XX}-U_X\right] + \mu(t)U_X \\ &+ \frac{1}{2}\left(\varepsilon^2 + \beta^2\sigma^2\left(1+2Y+Y^2\right)\right)U_{YY} - \kappa Y U_Y \\ &+ \beta\sigma^2\left(1+2Y+Y^2\right)U_{XY} - r(t)U = 0 \end{aligned} \tag{4.2}$$

where $r(t)$ is the discount rate and subscripts denote partial derivatives.

4.1 Calibration of the local volatility

The parameters of the stochastic volatility, β, ε and κ are specified before the calibration. Given that these parameters are specified, we calibrate the local volatility

$\sigma \equiv \sigma_{LSV}(t, S)$ so that the vanilla surface is matched by construction. For calibration of $\sigma_{LSV}(t, S)$ we use the standard relationship for local SV models:

$$\sigma^2_{LSV}(T, K)\mathbb{E}\left[(1 + Y(T))^2 \mid S(T) = K\right] = \sigma^2_{LV}(T, K) \tag{4.3}$$

where $\sigma^2_{LV}(T, K)$ is the local Dupire volatility.

The above expectation is computed by solving the forward PDE corresponding to pricing PDE (4.2) using finite-difference methods and computing $\sigma^2_{LSV}(T, K)$ stepping forward in time (see Sepp (2011) for details). Once $\sigma_{LSV}(t, S)$ is calibrated we use either backward PDE-s or MC simulation for valuation of exotic options.

5 Properties of the volatility process

In this section we consider pricing model (2.2) with constant volatility σ_{CV}.

5.1 The instantaneous variance of variance

The instantaneous variance of $Y(t)$ is given by:

$$dY(t)dY(t) = \left(\beta^2\sigma^2_{CV}(1 + Y(t))^2 + \varepsilon^2\right) dt \tag{5.1}$$

which has the systemic part and idiosyncratic part ε. In a stress regime, for large values of $Y(t)$, the variance is dominated by $\beta^2\sigma^2_{CV}Y^2(t)$. While in a normal regime, the variance is nearly a constant $(\beta^2\sigma^2_{CV} + \varepsilon^2)$. The variance of variance can be increased by increasing β (so the model approaches a pure SV models) or ε (so the model approaches a local SV models). Regimes with high volatility of volatility, say with volatility of the VIX of 100%, are reproduced by a combination of high values of σ_{CV} (local volatility σ^2_{LSV} in local beta SV model), β and ε.

5.2 The steady-state variance of variance

Now, we show that the volatility process has the steady-state volatility, so that the volatility approaches stationary distribution in the long run. We consider $\Upsilon(t) = Y^2(t)$ using equation (2.2):

$$d\Upsilon(t) = \left(-2\kappa\Upsilon(t) + \varepsilon^2 + \beta^2\sigma^2_{CV}(1 + 2Y(t) + \Upsilon(t))\right) dt + ... \tag{5.2}$$

We define $\overline{Y}(t) = \mathbb{E}\left[Y(t) \mid Y(0) = 0\right]$ and $\overline{\Upsilon}(t) = \mathbb{E}\left[\Upsilon(t) \mid Y(0) = 0\right]$. Using (5.2), we obtain: $\overline{Y}(t) = 0$ and

$$\overline{\Upsilon}(t) = \frac{\varepsilon^2 + \beta^2\sigma^2_{CV}}{2\kappa - \beta^2\sigma^2_{CV}}\left(1 - e^{-(2\kappa - \beta^2\sigma^2_{CV})t}\right) \tag{5.3}$$

We observe the following: first, the effective mean-reversion for the volatility of variance is:

$$2\kappa - \beta^2\sigma^2_{CV}$$

so that we need to enforce: $\kappa > \frac{1}{2}\beta^2\sigma_{CV}^2$; second, the steady state variance of volatility is

$$\frac{\varepsilon^2 + \beta^2\sigma_{CV}^2}{2\kappa - \beta^2\sigma_{CV}^2}$$

For high volatility σ_{CV}, the mean-reversion rate decreases, while the volatility of variance increases, which is in line with observed behavior in model behavior following "Risk Off" regimes, as we explained in Section 3. For the SV model based on the Ornstein-Uhlenbeck process, the steady state volatility of variance is given by $\frac{\varepsilon_{OU}^2}{2\kappa_{OU}}$, so that we obtain the following relationship:

$$\kappa_{OU} = \kappa - \frac{1}{2}\beta^2\sigma_{CV}^2 \ , \ \varepsilon_{OU}^2 = \varepsilon^2 + \beta^2\sigma_{CV}^2$$

5.3 Instantaneous correlation

We consider the instantaneous correlation between $dY(t)$ and $dX(t)$ using (4.1):

$$\rho(dX(t)dY(t)) = \frac{\beta\sigma_{CV}^2(1+Y(t))^2}{\sqrt{(\varepsilon^2 + \beta^2\sigma_{CV}^2(1+Y(t))^2)}\sqrt{\sigma_{CV}^2(1+Y(t))^2}} \tag{5.4}$$

With high volatility $Y(t)$ is large so letting $Y(t) \to \infty$ in (5.4) we obtain that $\rho(dX(t)dY(t)) = -1$, since $\beta < 0$. On the other hand, in a normal regime, $Y(t) \approx 0$, so that obtain:

$$\rho(dX(t)dY(t))|_{Y(t)\approx 0} = -\frac{1}{\sqrt{\left(\frac{\varepsilon^2}{\beta^2\sigma_{CV}^2} + 1\right)}} \ , \ \rho(dX(t)dY(t))|_{Y(t)\approx\infty} = -1 \tag{5.5}$$

Thus, the beta SV model introduces the state dependent spot-volatility correlation, with high volatility leading to absolute negative correlation. In contrast, the Heston and Ornstein-Uhlenbeck based SV models always assume constant instantaneous correlation.

Now, we consider the case when $Y(t)$ is close to -1, so that the instantaneous volatility is very small. Assume that $Y(t) = -1 + \delta$, $|\delta|$ is small. Substitution into yields:

$$\rho(dX(t)dY(t))|_{Y(t)=-1+\delta} = -\frac{1}{\sqrt{\left(\frac{\varepsilon^2}{\delta^2\beta^2\sigma_{CV}^2} + 1\right)}} \to -1 \text{ as } \delta \to 0$$

Thus the correlation approaches -1 for small instantaneous volatility. We note that if $\varepsilon = 0$ then $1+Y(t)$ would be log-normal process, so that point -1 is not reachable. Since ε is small, probability of $Y(t)$ going below point -1 is negligible.

Finally, we note that we can use the first equation in (5.5) to estimate the idiosyncratic volatility of volatility ε given a specified spot-volatility correlation ρ^* by equating it to given correlation ρ^* to obtain:

$$\varepsilon^2 = \sigma_{CV}^2\beta^2\frac{1-(\rho^*)^2}{(\rho^*)^2} \tag{5.6}$$

5.4 The steady state density

The steady state density function $G(Y)$ of volatility factor $Y(t)$ in dynamics (2.2) solves the following equation:

$$\frac{1}{2}\left[\left(\varepsilon^2+\beta^2\sigma_{CV}^2\left(1+2Y+Y^2\right)\right)G\right]_{YY}+\left[\kappa YG\right]_Y=0 \tag{5.7}$$

We can show that $G(Y)$ exhibits the power-like behavior for large values of Y:

$$\lim_{Y\to+\infty} G(Y)=Y^{-\alpha}\ ,\ \alpha=2\left(1+\frac{\kappa}{(\beta\sigma_{CV})^2}\right) \tag{5.8}$$

This power-like behavior contrasts with Heston and exponential volatility models which imply exponential tails for the steady-state density of the volatility. Thus, the beta SV model predicts much higher probabilities of large values of instantaneous volatility. This requires some care for numerical PDE implementation. For numerical PDE methods we specify the upper bound Y_∞ so that:

$$\mathbb{P}\left[Y>Y_\infty\right]\equiv\frac{1}{\alpha-1}Y_\infty^{1-\alpha}\le\epsilon \text{ thus } Y_\infty\ge\left[(\alpha-1)\epsilon\right]^{\frac{1}{\alpha-1}} \tag{5.9}$$

To find the lower bound, we can show that, in the vicinity of $Y=-1$, $G(Y)$ is Gaussian:

$$\lim_{Y\to-1} G(Y)=\mathbf{n}\left(\frac{Y+1}{\eta}\right)\ ,\ \eta^2=\frac{\varepsilon^2}{2(\kappa+(\beta\sigma_{CV})^2)} \tag{5.10}$$

where $\mathbf{n}(x)$ is normal PDF. Thus, the lower bound Y_0 can be specified by $Y_0\le\eta\mathbf{N}^{-1}(\epsilon)-1$, where $\mathbf{N}^{-1}(x)$ is the inverse of normal CPDF. For typical model parameters and $\epsilon=10^{-4}$, $Y_\infty\approx15$ and $Y_0\approx-3$. For our PDE solver, we use re-fined grid with number of points of 800 in Y direction and of 250 in time and spot directions.

6 Illustrations

6.1 Model calibration

The quality of fit of the beta SV model is similar to one-factor SV models - the model fits well longer-term skews (above one-year) while it is unable to fit short-term skews (up to one year) unless beta parameter β is large. To model the short-term skews we either resort to jumps or local volatility. We choose to introduce the local volatility.

To calibrate the model, we first use the pricing model with deterministic volatility and set the parameter β so that the model fits the market implied long-term 105%-95% skew. We found that the following empirical rule based on equation (2.7) works rather well:

$$\beta=\frac{\sigma_{IMP}(6m,5\%)-\sigma_{IMP}(6m,-5\%)}{0.05\sigma_{IMP}(6m,0\%)} \tag{6.1}$$

where $\sigma_{IMP}(6m, k\%)$ is 6m implied volatility for forward-based log-strike k.

Then we use equation (5.6) to specify ε given the beta specified as above and $\sigma_{CV} = \sigma_{DV}(1y)$ (a couple of iterations is needed as $\sigma_{DV}(1y)$ mildly depends on ε). We set the correlation as follows $\rho^* = -0.80$. The mean-reversion κ is fixed so that the shape of the model implied 105%-95% skew fits the market skew above 1y. Applying these approach to market data of options on the S&P500 index as of 20-June-2012, we obtain:

$$\beta = -4.74\ ,\ \varepsilon = 0.86\ ,\ \kappa = 1.10 \tag{6.2}$$

6.2 Impact of Model Parameters on Term Structure of Current Skew

In Figure 2, top, we show the impact of parameters β, ε and both of them, respectively, on the term structure of the model implied $105\% - 95\%$ skew using model parameters in (6.2) as the base case. We use beta SV model with deterministic $\sigma_{DV}(t)$ fitted to match the term structure of ATM volatilities for any combination of model parameters. Here, we shift β down and up by 50% to $\beta = -7.11$ and $\beta = -2.37$, respectively, and ε up and down by 50% to $\varepsilon = 1.29$ and $\varepsilon = 0.43$, respectively.

We see that, in the base case, the beta SV with deterministic volatility fits the term structure of the skew above one year very well. Decreasing β (in absolute value) leads to smaller overall level of skew. Increasing β leads to steeper skew in the short- and middle-term but the effect is not parallel as the convexity effect reduces the skew for longer maturities. Decreasing (increasing) ε leads to a roughly parallel downward (upward) shift in the term structure of the skew as the idiosyncratic component of the volatility of volatility declines (grows) increasing (decreasing) the skew. For shorter maturities, ε has little impact. Finally, increasing β (in absolute value) while decreasing ε leads to a roughly parallel downward shift in the model implied skew as a smaller value of the idiosyncratic volatility ε mitigates the convexity coming from a high value of β. As a result, by changing values of parameters β and ε, we can generate a variety of shapes for the term structure of the model implied skew.

6.3 Impact of Model Parameters on Forward Skew

In Figure 3, bottom, we show the impact of parameters β, ε and both of them, respectively, on the forward 3y3m skews for forward-start options starting in 3y and maturing in 3m. Here, we use the beta SV model with local volatility so that the model produces the same vanilla skews for different combinations of SV model parameters. In comparison, we plot the current 3m implied skew, and the forward skews computed by the local volatility. In general, we see that the local volatility predicts flattening skews; in contrast, the beta SV model produces steep forward skews.

We see that for larger (in absolute value) is β, the forward skews is steeper. Increasing ε, the convexity increases and the skew becomes U-shaped. However, the impact of ε is less pronounced for forward skews of OTM puts. Finally, increasing β and reducing ε leads to steep forward skews with forward 3y3m 105%-95% skew being nearly as steep as the current 3m skew. As a result, by changing values of parameters β and ε, we can produce different shapes of forward skews.

6.4 Pricing Cliquets

Now we apply the beta SV model for pricing cliquets using parameters as in (6.2) with local volatility calibrated to the vanilla surface. For comparison, we provide prices computed by the local volatility model and the Heston model (calibrated to skews at maturities 1y, 2y, and 3y). For the beta SV model we use three sets of parameters: LSV Beta (I) corresponds to the base case with $\beta = -4.74$, $\varepsilon = 0.86$; LSV Beta (II) corresponds to the model with increased β, $\beta = -7.11$, $\varepsilon = 0.86$; LSV Beta (III) corresponds to the model with increased β and decreased ε, $\beta = -7.11$, $\varepsilon = 0.43$.

Our example is the cliquet option with the following pay-off:

$$\sum_{i=1}^{N} \max\left[\min\left\{\frac{S(t_i)}{S(t_{i-1})} - 1, C\right\}, F\right]$$

In Table 2, we show prices of cliquet options with quarterly fixings and local floor (F) and cap (C) of -5% and 5%, respectively. This product depends heavily on the forward skew. In the local volatility model, the cap is overpriced while the floor is underpriced so that the cliquet price is too small. The Heston model produces markedly higher prices than the local volatility model, but it is known that observed market prices of cliquets are higher than those that can be produced by the Heston model. The base case model (I) under-prices cliques compared to Heston, while increasing β produces higher prices in model (II), while the strongest effect is produced with decreased ε in model (III).

Thus, using the beta SV model we can model the forward skew and volatility of volatility, and, as a result, prices of exotic options by changing parameters β and ε or introducing the term structure for these parameters (we recall that the beta SV model fits the vanilla surface by construction for any choice of β and ε).

7 Conclusions

We have presented the beta stochastic volatility model. One of the important features of this model is that its key parameter β has a natural interpretation as the rate of change in the short term ATM volatility given change in the stock price. As a result, empirical estimation of model parameters is easy to implement and interpret. In general, the beta SV model is augmented with local volatility, so that users can concentrate on modeling of forward skews with the vanilla surface being matched by construction. In case of deterministic volatility, we have derived an accurate approximation for call prices in this model. We have shown than the beta

SV model produces steeper forward skews and can be better suited for modeling of the forward volatility and related products.

References

[1] Bardhan I and Karasinski P, 1993 *Random volatility model implied by S&P option prices* Goldman Sachs internal paper

[2] Bergomi L, 2004 *Smile dynamics* Risk, February, pages 117-123

[3] Hagan, P., Kumar, D., Lesniewski, A., and Woodward, D. (2002), *Managing smile risk* Wilmott Magazine, September, 84-108

[4] Heston S, (1993), *A closed-form solution for options with stochastic volatility with applications to bond and currency options* Review of Financial Studies 6, 327-343

[5] Langnau A, 2004 *Introduction to the stochastic implied beta model* Merrill Lynch internal paper

[6] Lipton A, 2001 *Mathematical methods for foreign exchange: a financial engineer's approach* World Scientific, Singapore.

[7] Lipton A, 2002 *The vol smile problem* Risk, February, pages 81-85

[8] Sepp A, 2011 *Efficient numerical PDE methods to solve calibration and pricing problems in local stochastic volatility models* Global Derivatives Conference, Paris

8 Appendix A: Approximation for call price

Here we discuss an affine approximation for pricing equation (4.2) with constant or deterministic volatility σ. We use affine approximation including the second order terms:

$$G(t,T,X,Y;\Phi) = \exp\left\{-\Phi X + A^{(0)} + A^{(1)}Y + A^{(2)}Y^2\right\}$$

with $A^{(n)}(T;T) = 0$, $n = 0,1,2$.

By substitution this into PDE (4.2) and collecting terms proportional to Y and Y^2 only, we obtain a system of ODE-s for $A^{(n)}(t)$:

$$A_t^{(0)} + v_0 A^{(2)} + \frac{1}{2}v_0 (A^{(1)})^2 - \Phi A^{(1)} c_0 + \frac{1}{2}q = 0$$
$$A_t^{(1)} + \frac{1}{2}v_1 (A^{(1)})^2 + 2v_0 A^{(1)} A^{(2)} + v_1 A^{(2)} - \kappa A^{(1)} - \Phi\left(2c_0 A^{(2)} + c_1 A^{(1)}\right) + q = 0$$
$$A_t^{(2)} + \frac{1}{2}v_2 (A^{(1)})^2 + 2v_0 (A^{(2)})^2 + 2v_1 A^{(1)} A^{(2)} + v_2 A^{(2)} - 2\kappa A^{(2)} - \Phi\left(2c_1 A^{(2)} + c_2 A^{(1)}\right) + \frac{1}{2}q = 0$$

where

$$q = \sigma^2\left(\Phi^2 + \Phi\right)\ ,\ v_0 = \varepsilon^2 + \beta^2\sigma^2,\ v_1 = 2\beta^2\sigma^2,\ v_2 = \beta^2\sigma^2,\ c_0 = \beta\sigma^2,\ c_1 = 2\beta\sigma^2,\ c_2 = \beta\sigma^2$$

This is system is solved by means of Runge-Kutta fourth order method in a fast way. It is straightforward to incorporate time-dependent model parameters (but not space-dependent local volatility $\sigma_{LSV}(t,X)$). Thus, for pricing vanilla options, we can apply the standard methods based on the Fourier inversion. In particular, the value of the call option with strike K is computed by applying Lipton's formula (Lipton (2001)):

$$C(t,T,S,Y) = e^{-\int_t^T r(t')dt'}\left(e^{\int_t^T \mu(t')dt'} S - \frac{K}{\pi}\int_0^\infty \Re\left[\frac{G(t,T,x,Y;ik-1/2)}{k^2+1/4}\right]dk\right) \tag{8.1}$$

where $x = \ln(S/K) + \int_t^T \mu(t')dt'$.

9 Figure 1: Scatter plots

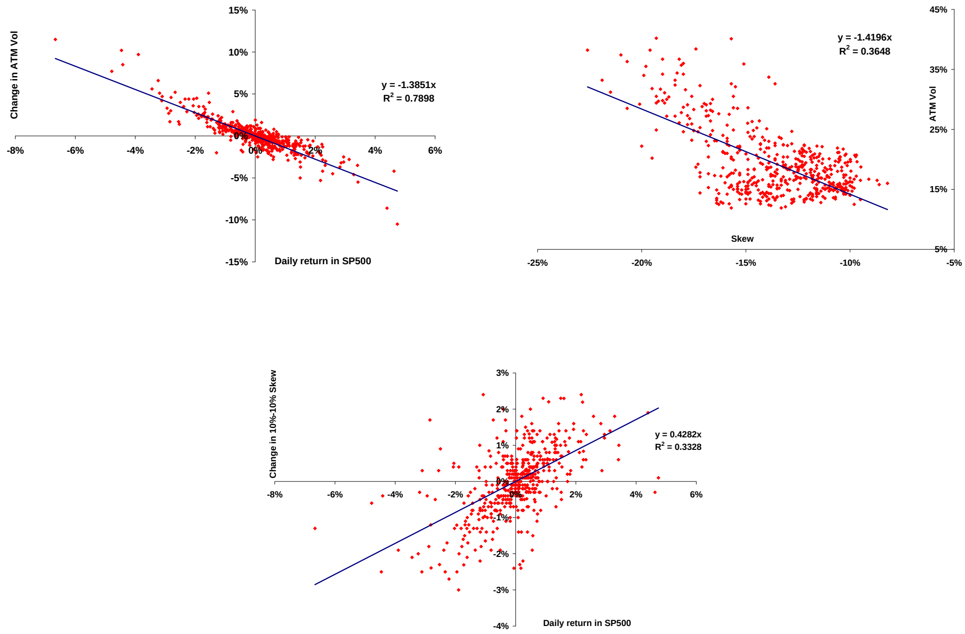
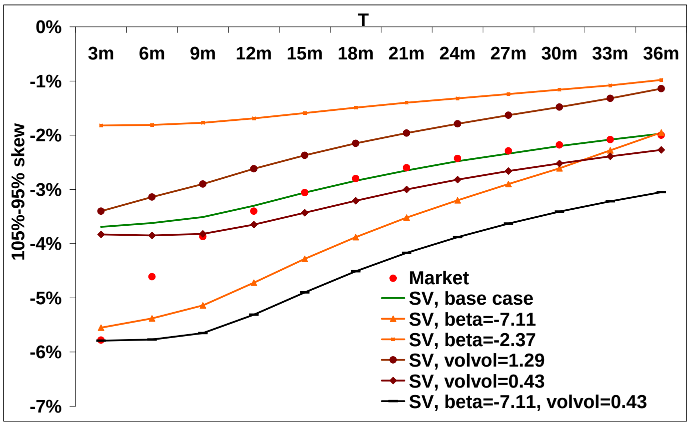


Figure 1: The scatter plot of daily changes in 1m ATM volatility against daily returns in the S&P500 index (top left), the scatter plot of 1m 110%-90% skew against the 1m ATM volatility (top right); the scatter plot of daily changes in 1m 110%-90% skew against daily returns in the S&P500 index using market data from October 2009 to October 2011 (bottom)

10 Figure 2: Impact of β and ε on term structure of current skew and the forward 3y3m skew



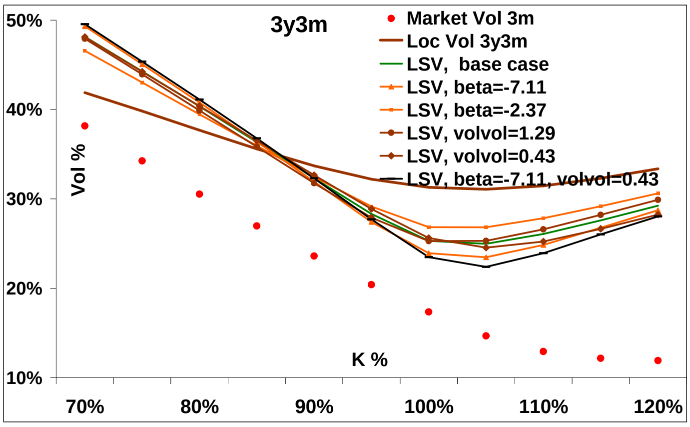


Figure 2: Top: impact on skew from β, ε, and their combination on the term structure of the model implied $105\% - 95\%$ skew using model parameters in (6.2) as the base case; bottom: impact on the forward 3y3m skew

11 Figure 3: Approximation

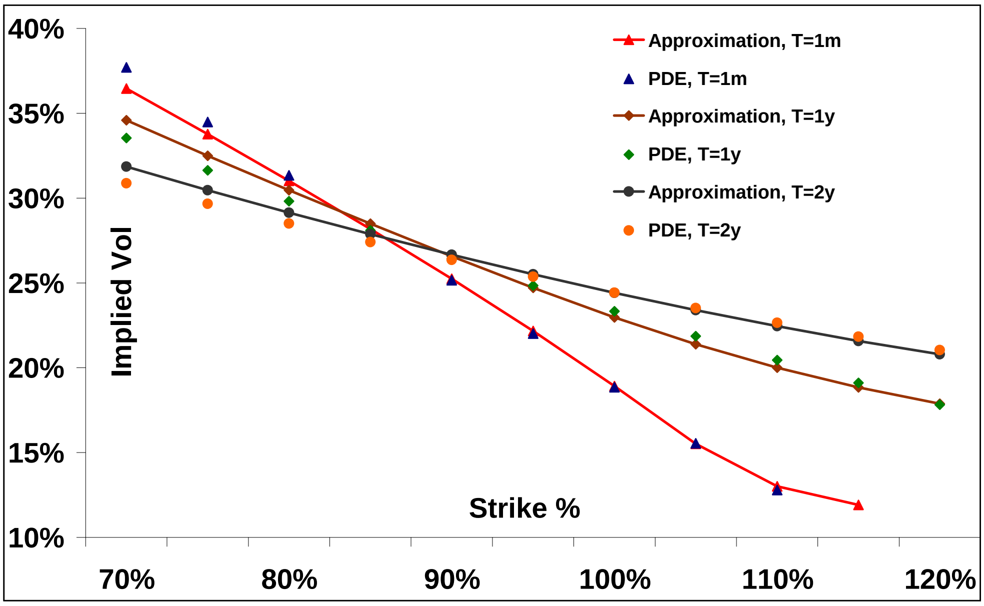


Figure 3: Implied model volatilities computed by approximation formula (8.1) against numerical PDE using $\beta = -7.63$, $\varepsilon = 0.35$, $\kappa = 4.32$ and constant volatilities: $\sigma_{ATM}(1m) = 19.14\%$, $\sigma_{ATM}(1y) = 23.96\%$, $\sigma_{ATM}(2y) = 24.21\%$

12 Table 1: Market and model data

Year	SP500	VIX	R SP500	R VIX	St SP500	St VIX	Beta	Kappa	St Resid	R2	Correl	Skew	State
1990	334.60	23.08%	-8.46%	9.42%	15.95%	26.97%	-1.00	14.66	21.17%	39%	-60%	-10.05%	Risk Off
1991	376.55	18.35%	25.57%	-7.67%	14.27%	19.39%	-0.74	10.12	15.82%	33%	-57%	-7.38%	Risk On
1992	415.82	15.41%	4.72%	-5.59%	9.72%	11.71%	-0.60	15.99	9.96%	28%	-50%	-5.98%	Risk On
1993	451.73	12.68%	7.05%	-0.79%	8.62%	11.45%	-0.64	31.04	9.57%	30%	-50%	-6.38%	Risk On
1994	460.39	13.94%	-0.89%	1.68%	9.84%	15.04%	-1.07	23.54	10.15%	55%	-71%	-10.71%	Range
1995	542.36	12.41%	30.48%	-2.06%	7.83%	9.94%	-0.50	54.27	8.46%	28%	-40%	-4.98%	Risk On
1996	670.95	16.50%	17.87%	8.95%	11.82%	15.71%	-0.87	30.78	11.06%	51%	-67%	-8.71%	Risk On
1997	874.37	22.39%	29.65%	2.28%	18.16%	19.69%	-0.75	6.86	13.79%	51%	-71%	-7.46%	Risk On
1998	1086.51	25.61%	25.14%	2.75%	20.29%	29.63%	-1.24	5.57	15.06%	74%	-86%	-12.37%	Risk On
1999	1328.23	24.35%	18.60%	-1.96%	18.10%	21.85%	-0.97	17.68	12.08%	69%	-82%	-9.65%	Risk On
2000	1426.54	23.32%	-10.08%	5.78%	22.38%	21.73%	-0.76	11.70	12.98%	64%	-79%	-7.62%	Risk Off
2001	1193.66	25.74%	-8.32%	-7.28%	21.21%	23.98%	-0.93	4.68	13.76%	67%	-82%	-9.27%	Risk Off
2002	992.96	27.29%	-20.49%	2.68%	26.24%	27.47%	-0.87	2.81	14.79%	71%	-84%	-8.74%	Risk Off
2003	966.02	21.95%	21.24%	-7.17%	16.76%	13.96%	-0.54	3.54	10.49%	44%	-65%	-5.43%	Risk On
2004	1131.02	15.46%	8.72%	-4.14%	11.12%	12.76%	-0.86	16.44	8.02%	61%	-76%	-8.64%	Risk On
2005	1207.49	12.80%	5.94%	-2.94%	10.38%	11.34%	-0.90	13.91	6.16%	71%	-83%	-8.96%	Risk On
2006	1311.05	12.81%	11.51%	0.90%	9.88%	14.18%	-1.13	14.91	8.26%	66%	-80%	-11.28%	Risk On
2007	1477.31	17.58%	3.42%	11.13%	16.02%	25.77%	-1.38	6.66	12.69%	76%	-87%	-13.81%	Range
2008	1218.01	32.76%	-35.54%	16.02%	41.17%	56.77%	-1.20	2.28	27.54%	77%	-87%	-11.98%	Risk Off
2009	948.84	31.40%	23.23%	-19.15%	27.15%	32.27%	-0.93	5.60	19.67%	63%	-79%	-9.35%	Risk On
2010	1140.52	22.54%	13.18%	-2.43%	18.02%	31.93%	-1.46	15.12	17.15%	71%	-83%	-14.56%	Risk On
2011	1267.66	24.22%	3.12%	5.36%	23.30%	39.92%	-1.51	4.01	18.48%	79%	-89%	-15.06%	Range

Table 1: Market and model historical data - econometric model

13 Table 2: Pricing Cliquets

	Local Vol	Heston	LSV Beta (I)	LSV Beta (II)	LSV Beta (III)
1y	1.03%	2.10%	1.38%	1.93%	1.82%
2y	0.67%	3.70%	2.25%	2.82%	3.69%
3y	0.01%	5.09%	3.16%	4.10%	5.99%

Table 2: Cliquet, LSV Beta (I) corresponds to $\beta = -4.74$, $\varepsilon = 0.86$; LSV Beta (II) corresponds to $\beta = -7.11$, $\varepsilon = 0.86$; LSV Beta (III) corresponds to $\beta = -7.11$, $\varepsilon = 0.43$