Some Stylized Facts of Short-Term Implied Volatility

Shiheng Shen

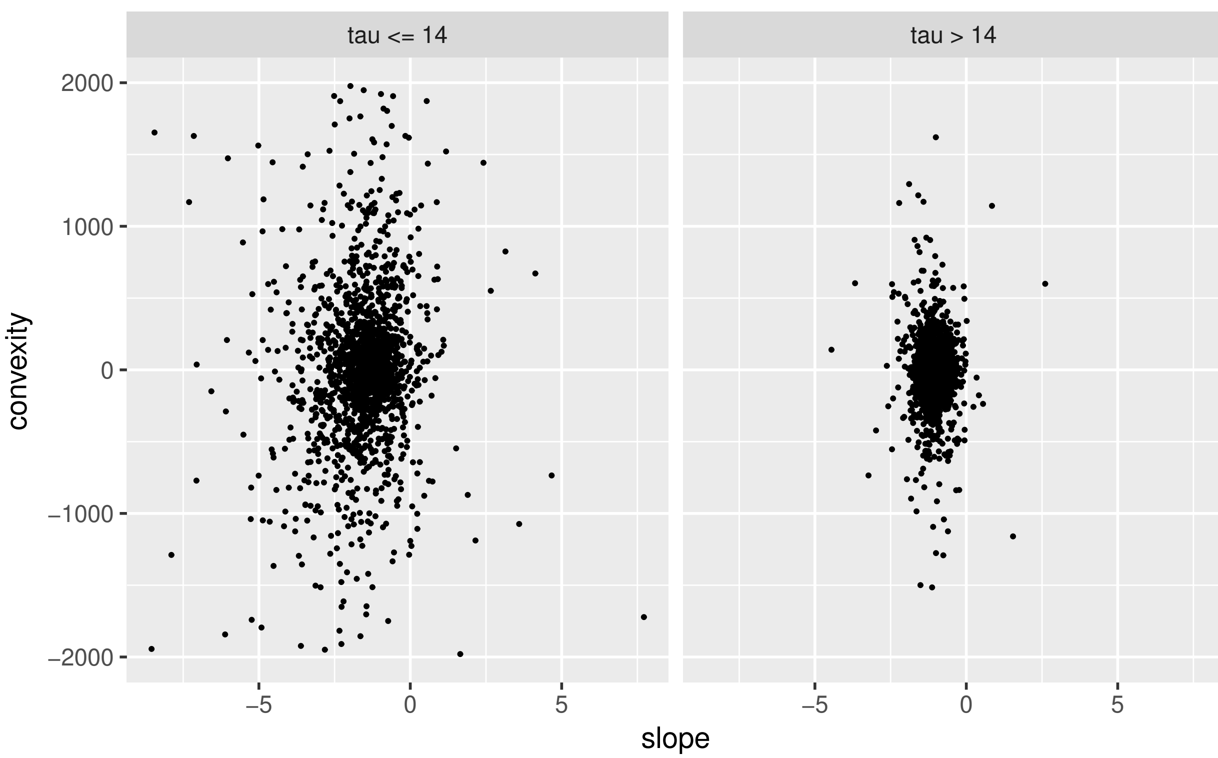
Stylized Fact 1: correlation between slope and convexity

Slope and convexity here denote $(\frac{\partial \Sigma}{\partial k})_{k=0}$ and $(\frac{\partial^2 \Sigma}{\partial k^2})_{k=0}$.

The scatter plot for data with $\tau = 1/252 \sim 29/252$, divided into two groups, their convexity against slope is shown in the following figure. The correlations for the two groups are $0.0932, -0.0085$.

A simple linear regression result shows that for group 1 slope and convexity are correlated, with $pvalue = 7.3564e-05$; for group 2 the correlation is not significant. Therefore we conclude that for relatively longer maturity, slope and convexity are nearly independent; but for shorter term, there might be dependence between slope and convexity.

From the figure, we can also observe that convexity has flexible sign. Slope is negative. At shorter maturity, both slope and convexity become more divergent.

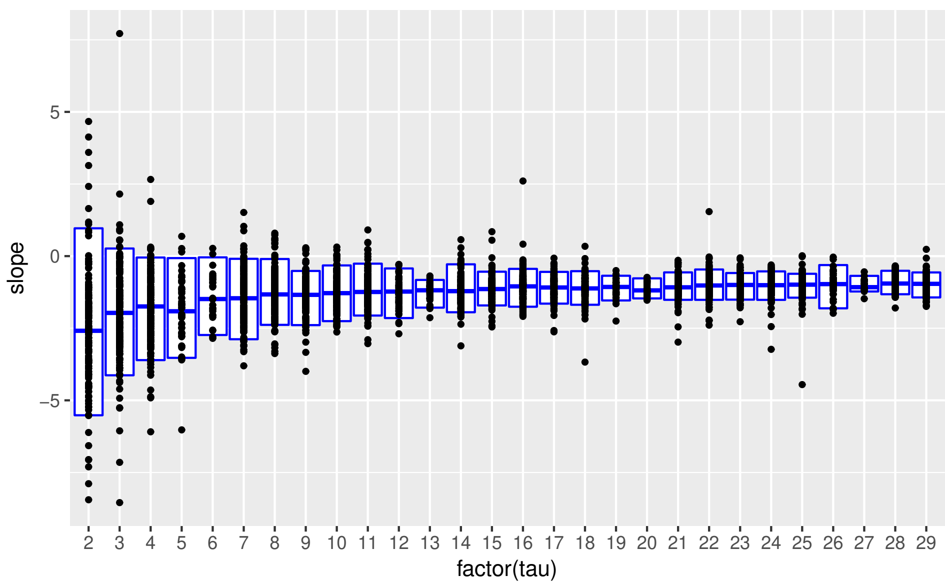


Stylized Fact 2: flattening slope with the increase of τ

After removing some extreme values[1], scatter plot of slope versus maturity is shown below, with median and $10\%, 90\%$ quantile marked in blue box.

It is known that the term structure of skew is approximately consistent with square-root (or at least power-law) decay[2]. Gatheral used data with maturity up to 700 days and he used $skew = \Sigma|_{F/K=1.05} - \Sigma|_{F/K=0.95}$ to measure at-the-money skew and got the relation $skew = C \cdot \tau^{-0.39}$.

I use $skew = -(\frac{\partial \Sigma}{\partial k})_{k=0}$ $(k = log(K/S))$ to measure at-the-money skew and fit a linear model for $log(skew) \sim log(\tau)$. The fitted model is $log(skew) = -0.37491 \cdot log(\tau) - 1.0403$ with $R^2 = 0.96$. This linear model amounts to $skew = C \cdot \tau^{\alpha}$, $\alpha = -0.37491$. The 95% confidential interval for α is $[-0.4084, -0.3414]$. So our estimate is consistent with previous work.

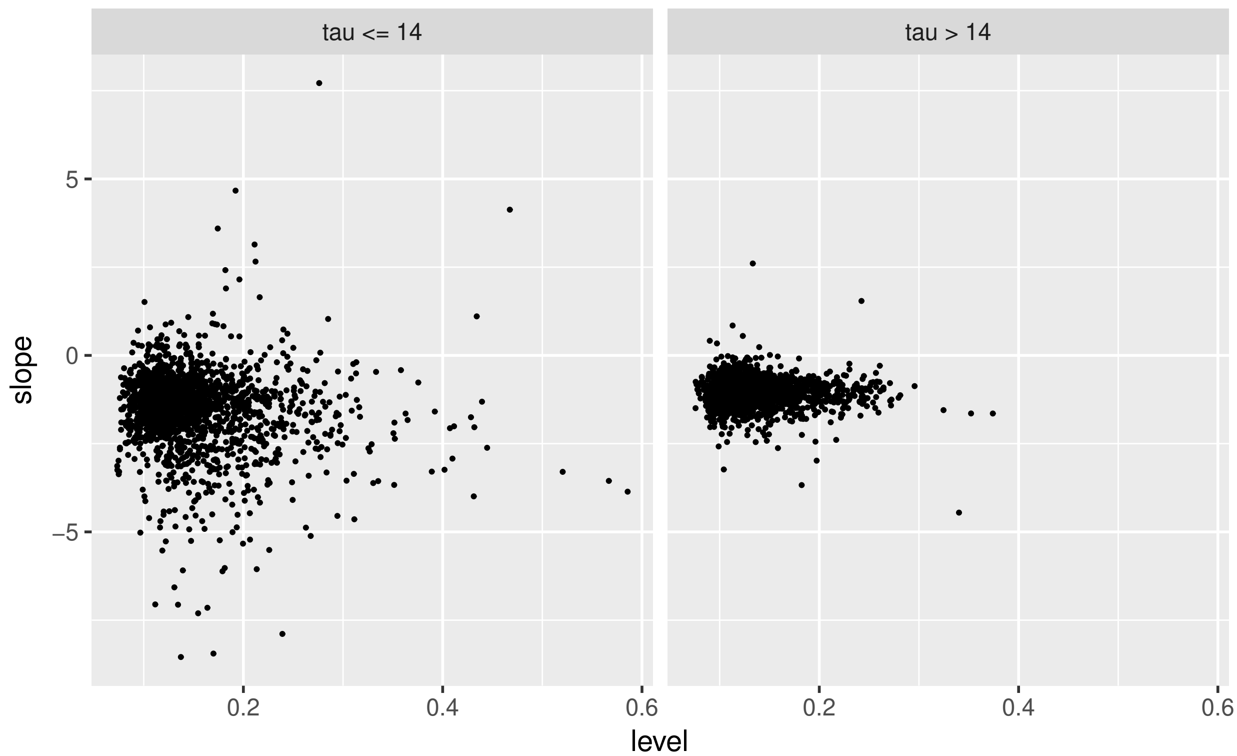


[1]Data with $\tau = 1/252$ is rare, so they are not included in the figure.
[2]Jim Gatheral's Notes

Stylized Fact 3: correlation between level and slope

The correlations for the two groups are $-0.127, -0.031, -0.048$. From the plot, we can see that group 2 and 3 are similar to the scatter plot in the previous paper. And their correlation is similar as well ($corr = -0.04$ in the previous paper).

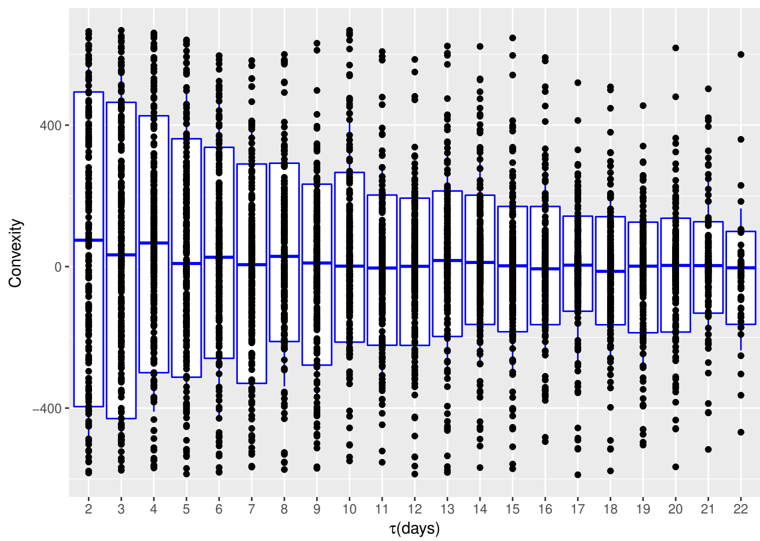
Noticeably level and slope are negatively correlated for group 1, though it's hard to detect anything from the plot. Linear regression shows that $slope = -1.78 \cdot level + -1.21$ and the p-value of the slope coefficient is 1.19×10^{-5}. So slope does tend to be smaller for higher level at $\tau = 1/252 \sim 22/252$.

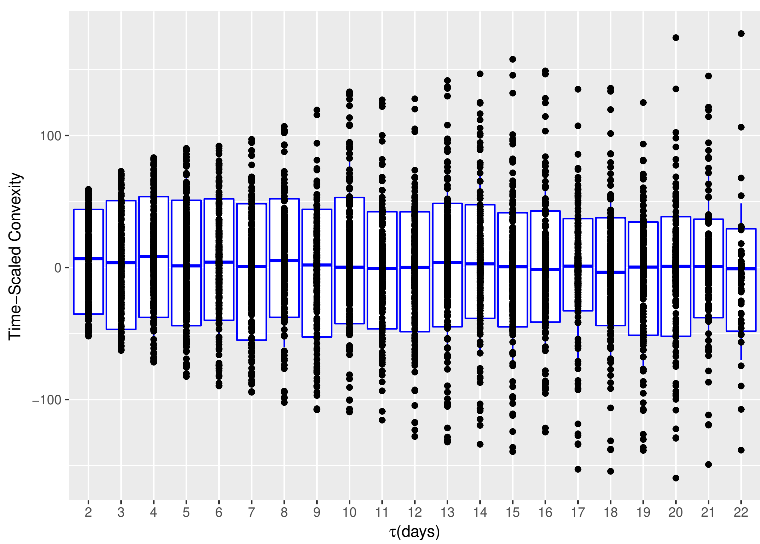


Stylized Fact 4: time-scaled convexity

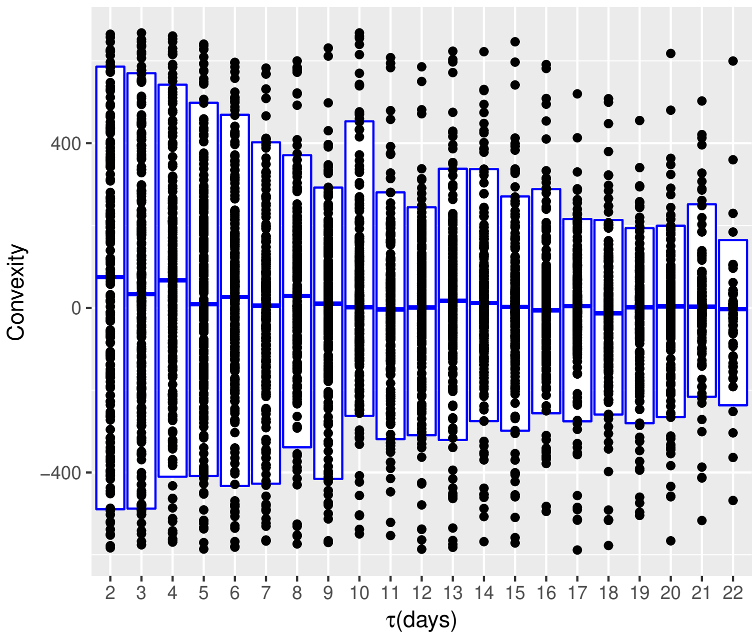
Scatter plot of (time-scaled) convexity versus maturity is shown in the following two figures, with a blue box at each τ indicating the 10% and 90% quantile.

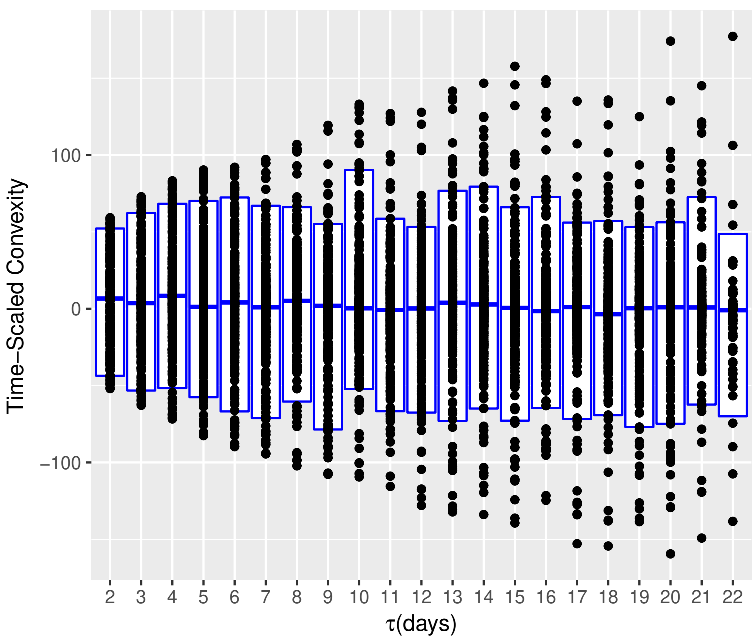
It can be seen that time-scaling help to concentrate the data. The $10\%, 90\%$ quantile of the time-scaled convexity is kind of invariant across maturity. If convexity follows a symmetric distribution, σ_{convex} should be related to the quantile. So we can guess that $Var(\sqrt{\tau} \cdot (\frac{\partial^2 \Sigma}{\partial k^2})_{k=0}) = const(\text{at different } \tau)$.





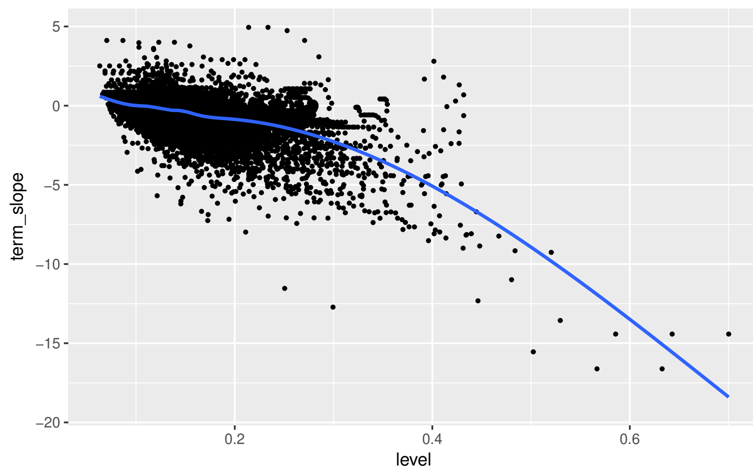
The following two figures highlight the 5%, 95% quantile.





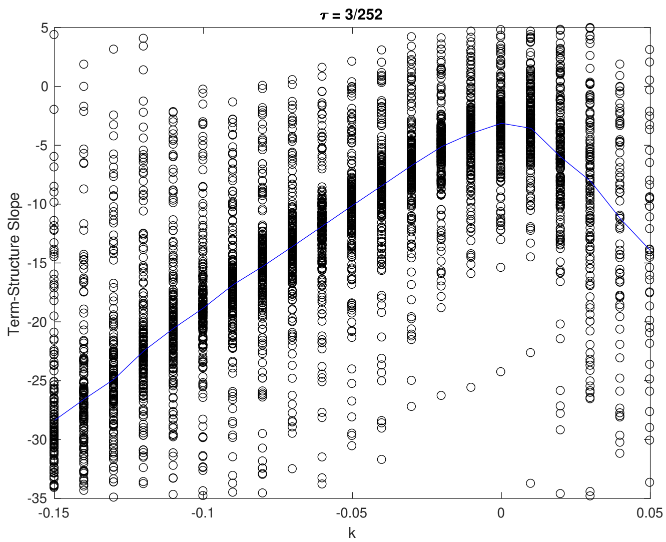
Stylized Fact 5: relationship between the at-the-money term-structure slope and level

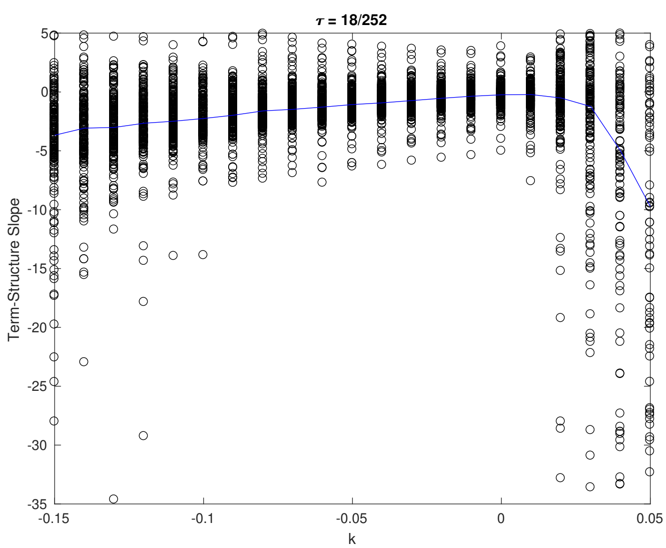
As shown in the following figure, this SF preserves at $\tau = 1/252 \sim 29/252$ (the fitted curve is the loess fit). The next figure plot the three groups separately (with a fitted curve using loess).



Stylized Fact 6: nonlinear relationship between term-structure slope and log-moneyness

The asymmetrical arch shape of term-structure slope versus log-moneyness exists across the sample. I use $\tau = 3/252, 18/252$ as representatives, shown in the following two figures, where the blue curves indicate the median of term-structure slope. We can observe that the shape is steeper for $\tau = 3/252$ than $\tau = 18/252$.

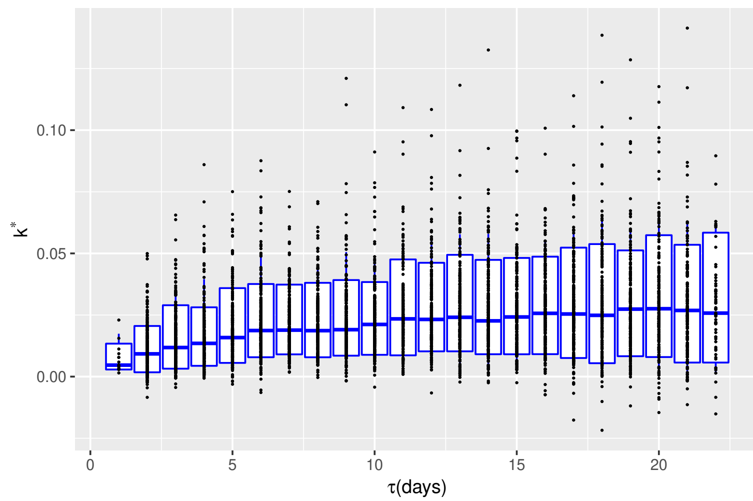


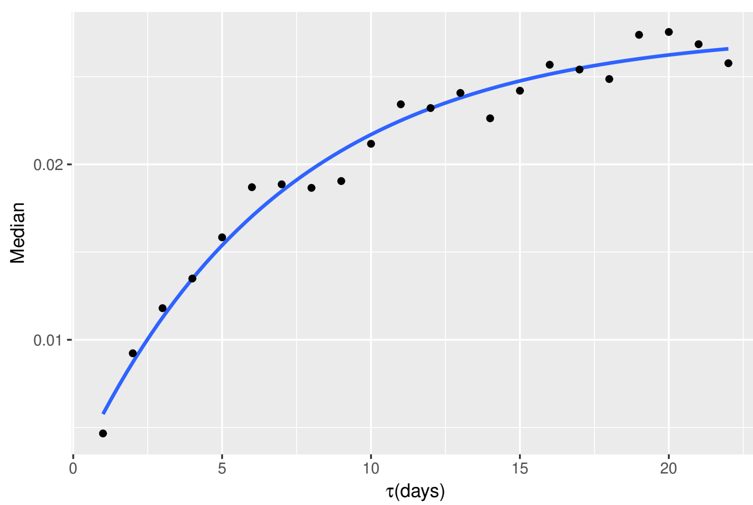


Stylized Fact 7: how k^* changes w.r.t. maturity

k^* here is the log-moneyness, at which the implied volatility curve achieves its minimum point. The scatter plot below used data on all available days. The blue box marks the $10\%, 50\%, 90\%$ quantile.

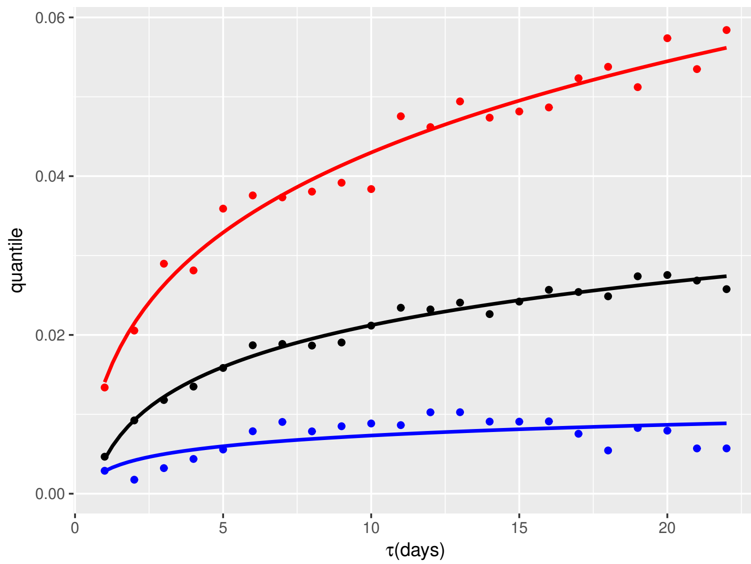
It can be seen that the median of k^* should follow a concave function of τ. I assume the form to be $c_1 - c_2 \cdot e^{-\lambda\tau}$. Using a grid search of λ, the model that yields highest R^2 is $Med(k^*) = 0.0276 - 0.0253 \cdot e^{-0.145\tau}$, with $R^2 = 0.977$.





Next we assume $k^* = c_1 + c_2 \cdot \tau^\alpha$. Performing the regression on the median and 10%, 90% quantile respectively (determine α by grid search). Results are shown in the following figure.

As can be seen, the 10% quantile is quite flat and the estimated α is less than 10^{-6}. The α for median and 90% quantile are 0.0434 and 0.1904, with $R^2 = 0.980, 0.966$ respectively.



Stylized Fact 8: short-time variance skew behavior

Motivation

I include this SF because short-time variance skew has special property which the original volatility skew doesn't have.

Under Jump-Diffusion Model:

We first consider under this setting because it's simple and can be considered as prototype for a large class of more complex models such as the stochastic volatility plus jumps model.

We have $dS_t = \mu S_{t-}dt + \sigma S_{t-}dW_t + S_{t-}dJ_t$. Jump components are $J_t = \Sigma_{i=1}^{N_t} Y_i$, where N_t is a poisson process with intensity λ and jump sizes Y_i are i.i.d with distribution F. We define the jump compensator to be $\mu_J = -\lambda \int \xi dF(\xi)$.

Under this setting, intensity is finite, so when we consider very short maturity, multiple jumps are negligible. Denote the price of a call option C_J, then when a jump occur (stock price $S \to S'$) with small time δT to expiration

$$C_J(S, K, \delta T) \approx (1 - \lambda\delta T) \cdot C_{BS}(Se^{\mu_J \delta T}) + \lambda\delta T \cdot C_{BS}(S', K, \delta T) = C_{BS}(Se^{\mu_J \delta T}) + O(\delta T)$$

Considering $\dfrac{\partial C_J}{\partial k} = \dfrac{\partial C_{BS}}{\partial k} + \dfrac{\partial C_{BS}}{\partial \Sigma_{BS}} \dfrac{\partial \Sigma_{BS}}{\partial k}$

$$\frac{\partial \Sigma_{BS}^2}{\partial k}|_{k=0} = 2\Sigma_{BS}|_{k=0}[\frac{\partial C_J}{\partial k} - \frac{\partial C_{BS}}{\partial k}(\frac{\partial C_{BS}}{\partial \Sigma_{BS}})^{-1}]|_{k=0}$$

The behavior of terms appearing in the RHS is known ($\Sigma_{BS} \to \sigma_{BS}$, $\dfrac{\partial C_{BS}}{\partial \Sigma_{BS}} \to \dfrac{S}{\sqrt{2\pi}}\sqrt{\delta T}$), so we can obtain

$$lim_{\tau \to 0} \frac{\partial \Sigma_{BS}^2(k, \tau)}{\partial k}|_{k=0} \approx -2\mu_J$$

SVJ model

We next consider a model adding stochastic volatility:

$$\begin{aligned} dS_t &= \mu S_{t-}dt + \sigma_t S_{t-}dW_t^1 + S_{t-}dJ_t \\ d\sigma_t &= a(\sigma_t)dt + b(\sigma_t)dZ_t \\ dZ_t &= \rho dW_t^1 + \sqrt{1 - \rho^2}dW_t^2 \end{aligned}$$

where dW_t^1, dW_t^2 represent two independent brownian motions.

The variance skew are approximately additive for jump and stochastic volatility effects [3], $lim_{\tau \to 0} \dfrac{\partial \Sigma_{BS}^2(k, \tau)}{\partial k}_{k=0} \approx \rho b(\sigma) - 2\mu_J$.

These interesting features motivates us to consider this SF.

[3] J. Gatheral, *The Volatility Surface, Chapter 7*

The figure below (Left) plot $-\frac{\partial \Sigma^2_{BS}(k,\tau)}{\partial k}|_{k=0}$ with different τ. The figure is too zigzag, so we look at the median of $\frac{\partial \Sigma^2_{BS}(k,\tau)}{\partial k}|_{k=0}$ instead, shown in the plot on the right.

The jump size should be finite, which indicate convergence as $\tau \to 0$ so a model of $\tau^{-\alpha}$ is not very reasonable. Therefore we choose the form: $var.skew = -c_0 - c_1 \cdot exp(-\lambda\tau)$.

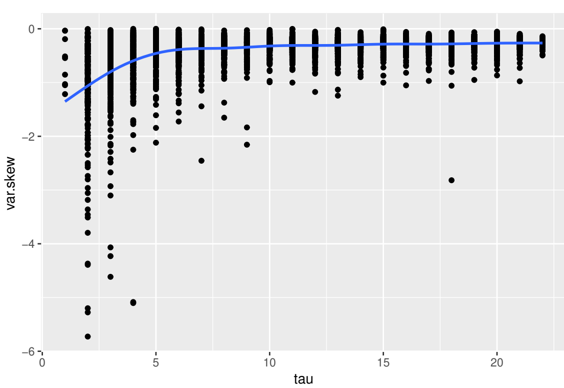
The fitted model is $med(\frac{\partial \Sigma^2_{BS}(k,\tau)}{\partial k}|_{k=0}) = -0.2319 - 0.2657 \cdot exp(-0.1328\tau)$ (here τ is measured in days), with $R^2 = 0.97$.

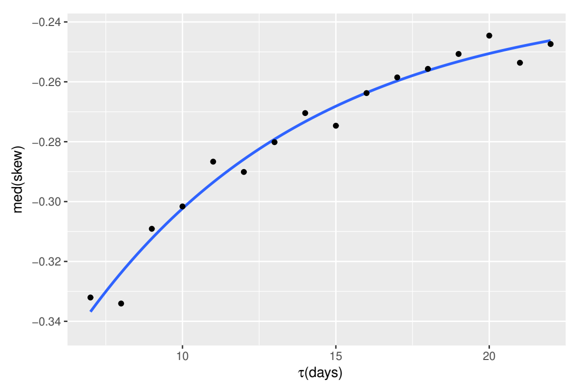
So we estimate the mean of short-time variance skew asymptotics

$lim_{\tau\to 0}\mathbb{E}\frac{\partial \Sigma^2_{BS}(k,\tau)}{\partial k}|_{k=0} \approx -0.497.$

Connection with jump distribution

According to definition, μ_J measures the (negative[4]) average jump size. Therefore, as jumps becomes more negatively skewed, μ_j will increases and the variance skew becomes more negative.





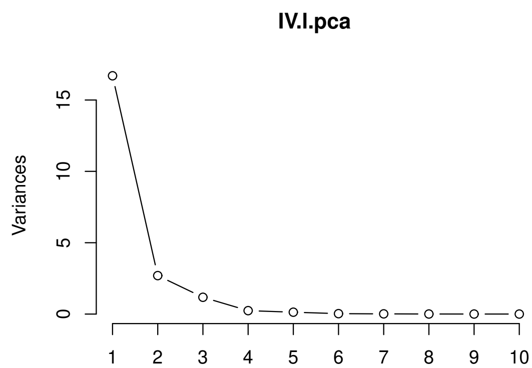
[4]Because we define $\mu_J = -\lambda \int \xi dF(\xi)$. It should be also mentioned that the integral should be taken at $[-1, \infty]$ because stock price must be positive.

Stylized Fact 9: PCA of IVS

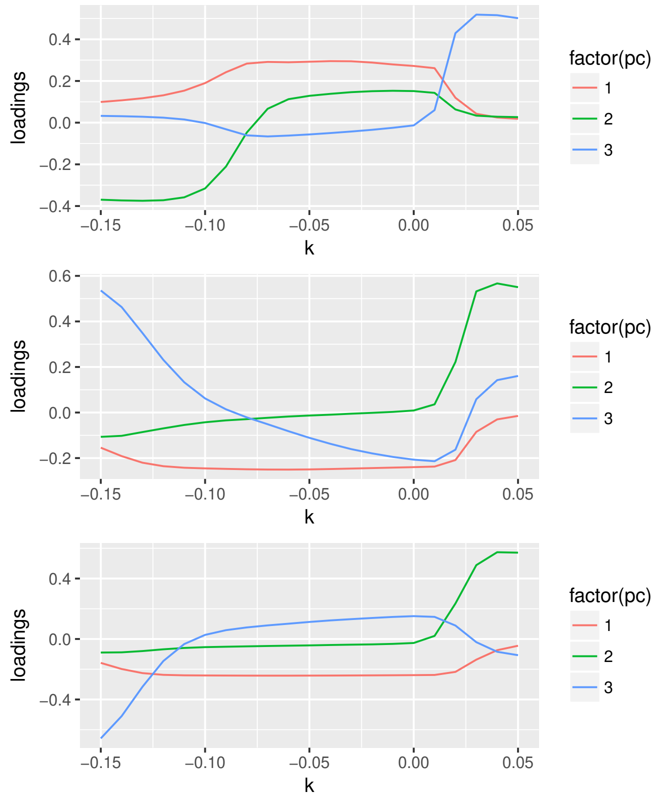
First separate the IVS data into three baskets based on maturity as in SF3 and SF5. We have interpolated IV for 21 log-moneyness $[-0.15, 0.05]$ for each maturity available on each day. We then apply standard PCA (centered and scaled, but no BoxCox transform so that we can reconstruct IVS easily using eigenvectors) for each basket.

We perform PCA directly on IV because for each τ, we don't have data with consecutive(or equally spacing) dates which prevent us from differencing IV first or using covariance matrix.

PCA is reasonable for IV (especially for longer maturity) as shown in the following figure (Taken group 3 as an example). It shows the variance explained by the first 10 PC. For group 3: $\tau = 15/252 \sim 22/252$, the first three PC explained $79.3\%, 12.9\%, 5.8\%$ of the variance respectively, together explaining 98.1% of the variance. Yet with group 1 and 2, the first three PC can only explain $91.9\%, 96.5\%$ of the total variance .



The loadings of the first three PC for the three groups are shown respectively in the following figure. It seemed to me that the loadings are a little hard to explain. The 1st PC is virtually the mean IV. The 2nd PC have similar shape in the three group but 3rd PC seems obscure.



Next, we look at the estimated distribution of the scores of the 1st PC for the three groups, shown in the following figure. It's obvious that the scores of the 1st PC becomes negatively skewed at shorter maturity.