MTH 9875 The Volatility Surface: Fall 2016

Lecture 12: The VIX index and its derivatives

Jim Gatheral
Department of Mathematics

Outline of lecture

- Volatility derivatives
  - Model-independence: bounds, conjectures and counterexamples
- The VIX index
- VIX futures and options
- The VVIX index

- Variance curve models
  - The Double CEV model
  - Dynamics consistent with VIX options prices
- VIX-based ETNs

Valuation of volatility derivatives given vanilla prices

- In Lecture 10, we saw that if volatility and stock price moves are uncorrelated, volatility derivatives may in principle be valued given just the prices of European options on stock.
  - The zero-correlation assumption is completely unrealistic but the connection between European option prices and volatility derivative values is suggestive.

- How tightly do the prices of European options (which are assumed to be known) constrain the fair value of a volatility derivative?

A simple lognormal model

Assume that $\log\left(\sqrt{\langle x\rangle_T}\right)$ is normally distributed with mean μ and variance s^2. Then $\log\left(\langle x\rangle_T\right)$ is also normally distributed with mean $2\,\mu$ and variance $4\,s^2$.

Volatility and variance swap values are given by respectively

$$\mathbb{E}[\sqrt{\langle x\rangle_T}] = e^{\mu+s^2/2};\ \mathbb{E}\left[\langle x\rangle_T\right] = e^{2\mu+2s^2}$$

Solving for μ and s^2 gives

$$s^2 = 2\log\left(\frac{\sqrt{\mathbb{E}\left[\langle X\rangle_T\right]}}{\mathbb{E}\left[\sqrt{\langle X\rangle_T}\right]}\right);\ \mu = \log\left(\frac{\mathbb{E}\left[\sqrt{\langle X\rangle_T}\right]^2}{\sqrt{\mathbb{E}\left[\langle X\rangle_T\right]}}\right)$$

and the convexity adjustment is given by

$$\sqrt{\mathbb{E}\left[\langle X\rangle_T\right]} - \mathbb{E}\left[\sqrt{\langle X\rangle_T}\right] = \left(e^{s^2/2}-1\right)\,\mathbb{E}\left[\sqrt{\langle X\rangle_T}\right]$$

Black-Scholes-style formula

Under this lognormal assumption, calls on variance may be valued using a Black-Scholes style formula:

(1)

$$\mathbb{E}\left[\langle x\rangle_T - K\right]^+ = e^{2\mu+2s^2}N(\tilde{d}_1) - K\,N(\tilde{d}_2)$$

with

$$\begin{aligned}\tilde{d}_1 &= \frac{-\frac{1}{2}\log K + \mu + 2s^2}{s}\\ \tilde{d}_2 &= \frac{-\frac{1}{2}\log K + \mu}{s}.\end{aligned}$$

- This simple lognormal assumption seems not too unreasonable
  - The empirical distribution of implied volatility changes looks lognormal.
  - Also, the dynamics of the volatility skew are consistent with approximately lognormal volatility dynamics.

- In this model, the variance and volatility swap values (or equivalently the variance swap value plus the convexity adjustment) are all that is required to fix the values of all options on variance.

- Moreover, at least for the major equity indices, there is a tight market in variance swaps and a somewhat less liquid market in the convexity adjustment.

A Heston example

Suppose the true dynamics of the underlying were Heston (recall that lognormal volatility is a much more realistic assumption) with BCC parameters:

$$\lambda = 1.15, \rho = 0, \sigma_0^2 = \bar{\sigma}^2 = 0.04, \eta = 0.39.$$

- We showed earlier how to compute the value of a volatility swap in terms of the Heston parameters.
- We can also easily compute the values of calls on variance in terms of the Heston parameters.

- Given the values of the variance and the volatility swap, we can also apply equation (1) to value the same calls on variance.

- How big is the error from applying the lognormal variance call formula?

Lognormal formula vs exact Heston computation

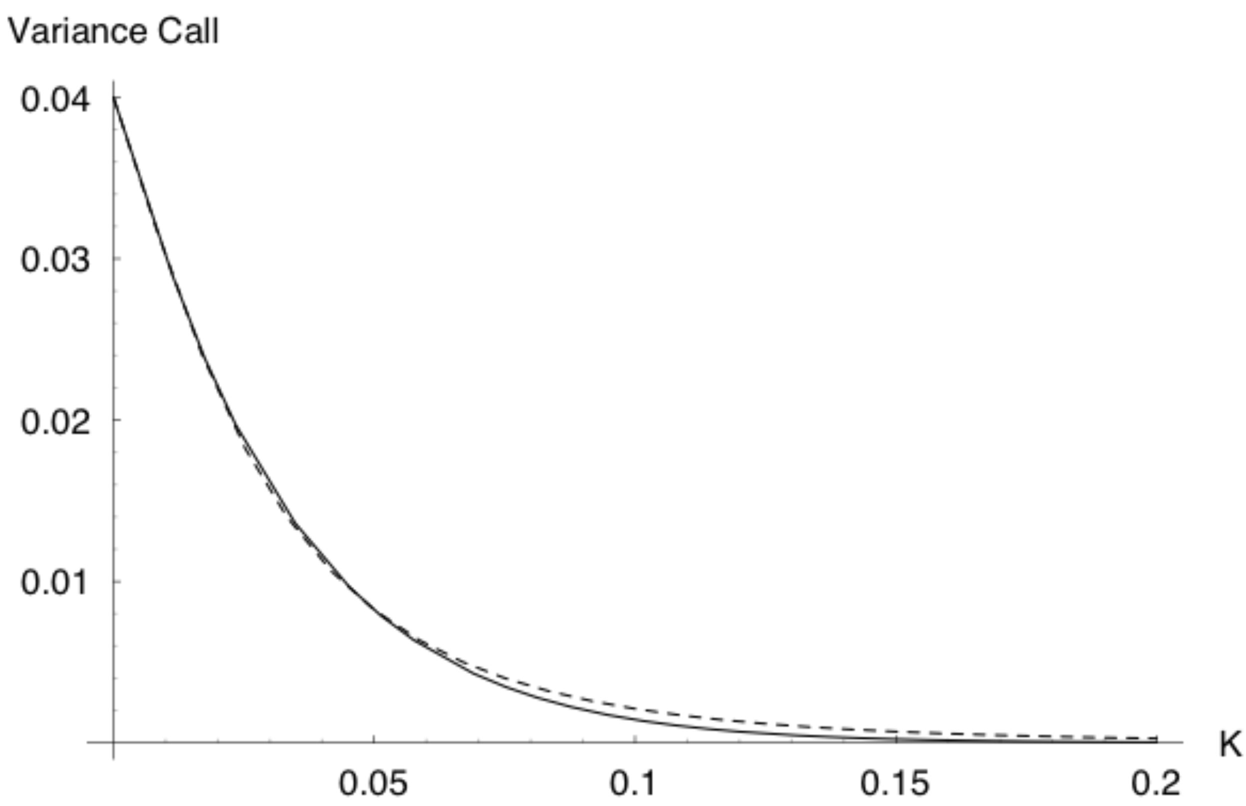


Figure 1: Value of one-year variance call vs variance strike K with the BCC parameters. The solid line is a numerical Heston solution; the dashed line comes from our lognormal approximation.

Remarks

- Obviously, the two approaches must agree at zero strike because then we have just a variance swap.
- It's not clear how big the valuation error really is away from zero strike.

- To get a better sense for the error, in Figure 2, we compare the density we get by inverting the Heston characteristic function numerically and the approximate lognormal density with mean $2\,\mu$ and variance $4\,s^2$.

Comparison of densities

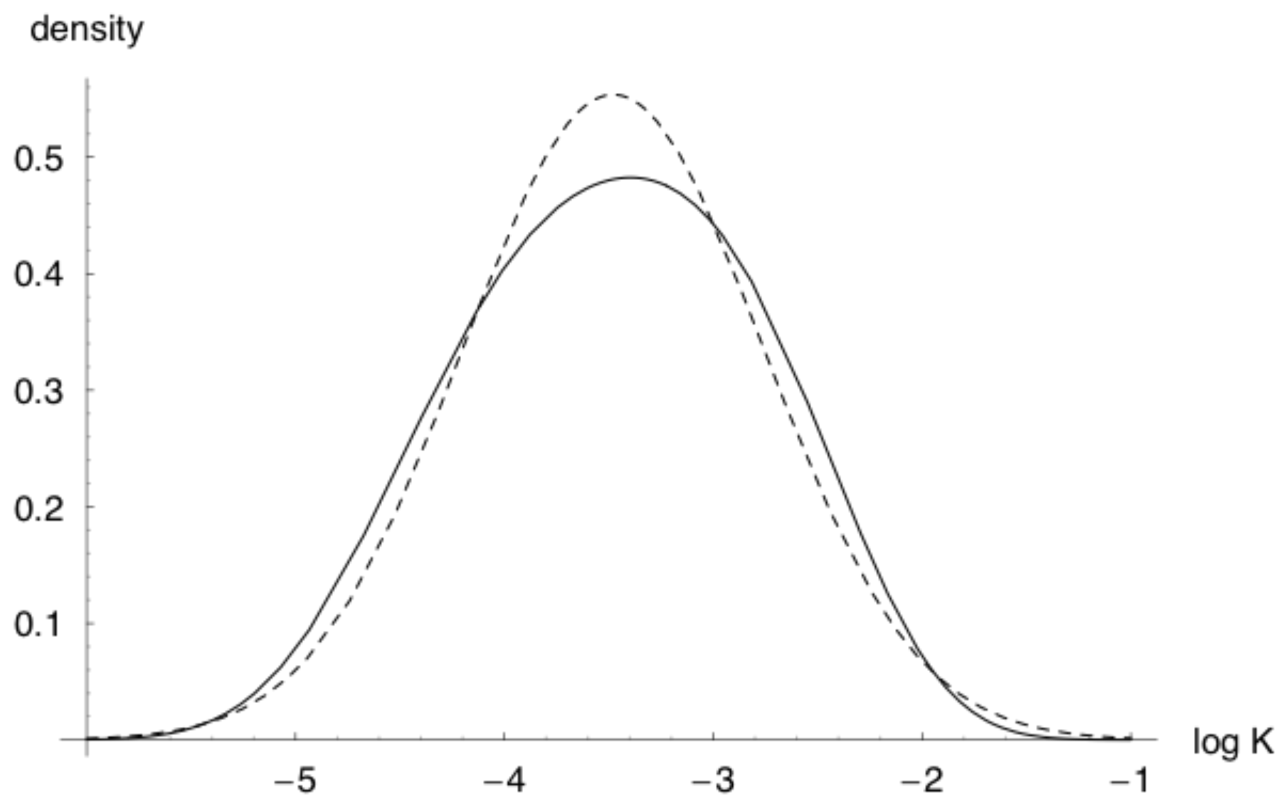


Figure 2: The pdf of the log of one-year quadratic variation with BCC parameters. The solid line comes from an exact numerical Heston computation; the dashed line comes from our lognormal approximation.

Constraints on volatility derivative pricing

- We see that agreement is pretty close.
  - And if agreement is so close in the Heston case where we believe the dynamics to be unrealistic, how much better should the agreement be when the underlying volatility dynamics are lognormal (the more realistic case)?

- We conclude that the values of T-maturity volatility derivatives are in practice tightly constrained by the prices of expiration-T European options.

Another potential recipe

- To make the argument even more explicit, we fit our preferred stochastic volatility model (lognormal let's say) to the prices of European options, and we set the correlation ρ to zero (this doesn't affect the value of volatility derivatives).
- Now recall the following formula from Lecture 11:

(2)

$$\mathbb{E}\left[\sqrt{\langle x\rangle_T}\right] = \sqrt{2\,\pi}\,\hat{c}(0) + \sqrt{\frac{2}{\pi}}\int_0^\infty e^{k/2}\,I_1\left(\frac{k}{2}\right)\,c(k)\,dk$$

where $I_n(\cdot)$ represents a modified Bessel function of the first kind.

- Compute the value of the volatility swap using (2) and apply our lognormal variance call option formula (1), confident that the result will be robust to the specific parameters.
  - The only constraint is that the fit to expiration-T European prices should be good.

Options on volatility: More on model-independence

- We have argued so far that given the prices of European options of all strikes and expirations, the values of options on variance should be tightly constrained.
  - The extent to which this is true and even how this claim should be precisely stated are as yet unresolved questions.

A model-independent lower bound

- [Dupire][6] shows how a model-independent lower bound for calls on realized variance may be obtained by considering the delta hedging of a portfolio of European call options.

- The idea is to hedge in business time defined as the timescale of realized quadratic variation.
  - Options are rehedged each time quadratic variation increases by a given amount.
  - Unlike conventional delta hedging in (atomic) clock time, assuming no jumps, this hedging strategy guarantees profits on a short option position if realized volatility is below implied volatility (the volatility implied by the initial sale price).
  - Conversely the hedged option position is guaranteed to lose if realized volatility is higher than implied at inception.

- Losses are insured against by buying an option on quadratic variation.
- The hedger ends up short a portfolio of European options and long an option on realized variance with only positive payoffs: this portfolio must be worth at least zero and a model-independent lower bound on the price of an option on variance follows.

- In the same work, Dupire shows (without formally proving) how solutions to the Skorokhod Embedding problem can generate effective upper and lower bounds on the price of an option on variance constrained by known prices of European options.

A wrong conjecture

- In [The Volatility Surface][8], it was conjectured that the minimum possible value of an option on variance is the one generated from a local volatility model fitted to the volatility surface.
  - Options on variance have value even in a local volatility model because realized variance depends on the realized path of the stock price from inception to expiration.

- Given that local variance is a risk-neutral conditional expectation of instantaneous variance, it seemed obvious that any other model would generate extra fluctuations of the local volatility surface relative to its initial state.

A counterexample due to Beiglböck et al.

Consider the following toy model from [Beiglböck et al.][3] where the volatility path can be one of

$$v_+ = \begin{cases} 2 & \text{if } t \in [0,1) \\ 3 & \text{if } t \in [1,2) \\ 1 & \text{if } t \in [2,3] \end{cases} \quad v_- = \begin{cases} 2 & \text{if } t \in [0,1) \\ 1 & \text{if } t \in [1,2) \\ 3 & \text{if } t \in [2,3] \end{cases}$$

depending on the toss of a coin.

Independent of the result of the coin-toss, the total variance

$$W_3 = \int_0^3 v_t \, dt = 6$$

so $\mathbb{E}[(W_3 - 6)^+] = 0$.

Now define

$$\Theta(x) = \partial_T C_{BS}; \; \Psi(x) = K^2 \, \partial_{K,K} C_{BS};$$

Then local variance is given by

$$
\begin{aligned}
v_\ell(x,t) = 2\,\frac{\partial_T C}{K^2\,\partial_{K,K} C} &= 2\,\frac{\Theta_+(x) + \Theta_-(x)}{\Psi_+(x) + \Psi_-(x)} \\
&= \frac{v_+(t)\,\Psi_+(x) + v_-(t)\,\Psi_-(x)}{\Psi_+(x) + \Psi_-(x)}
\end{aligned}
$$

which is obviously a nonconstant function of x and t.

One can verify with a Monte Carlo computation that

$$\mathbb{E}[(\tilde{W}_3 - 6)^+] \approx 0.026 > 0$$

In words, the value of the option on variance under local volatility is greater than the value of the option on variance under stochastic volatility in this case.

Carr and Lee upper and lower bounds

- Given European options prices, [Carr and Lee][4] derive model-independent upper and lower bounds for volatility derivative prices in a general diffusion model.
  - The lower bound comes from the Root construction which in some sense minimizes the variance of total variance for a given density ϕ.
  - The upper bound comes from the Rost construction which in some sense maximizes the variance of total variance.
  - Both of these constructions correspond to the Skorokhod embeddings originally suggested by Dupire.

- Carr and Lee show that these bounds are tight by explicitly constructing sub- and super-hedges for options on variance.
  - The sub-hedge is due to Dupire: Delta hedge a short call position in business time, making money if the realized variance is less than the implied variance at the time of sale.
  - The super-hedge involves replicating in business time a double knockout with a carefully chosen payoff at the barrier.

Conclusion

- Between these model-independent upper and lower bounds, it seems reasonable to suppose that the fair value of an option on variance is closest to that given by a stochastic volatility model.
  - It’s easy to reject local volatility.
  - The Rost model appears to represent crazy dynamics of the volatility surface whereas in practice, volatility surfaces are very well behaved.

The VIX index

- In 2004, the CBOE listed futures on the VIX - an implied volatility index.
- Originally, the VIX computation was designed to mimic the implied volatility of an at-the-money 1 month option on the OEX index. It did this by averaging volatilities from 8 options (puts and calls from the closest to ATM strikes in the nearest and next to nearest months).
- The CBOE changed the VIX computation: "CBOE is changing VIX to provide a more precise and robust measure of expected market volatility and to create a viable underlying index for tradable volatility products."

The new VIX formula

Here is the new VIX definition (converted to our notation) as specified in the CBOE white paper:

(3)

$$VIX^2 = \frac{2}{T} \sum_i \frac{\Delta K_i}{K_i^2} Q_i(K_i) - \frac{1}{T} \left[\frac{F}{K_0} - 1 \right]^2$$

where Q_i is the price of the out-of-the-money option with strike K_i and K_0 is the highest strike below the forward price F.

We recognize (3) as a straightforward discretization of the log-strip and makes clear the reason why the CBOE implies that the new index permits replication of volatility.

Specifically, (with obvious notation)

$$\begin{aligned}
\frac{VIX^2\, T}{2} &= \int_0^F \frac{dK}{K^2} P(K) + \int_F^\infty \frac{dK}{K^2} C(K) \\
&= \int_0^{K_0} \frac{dK}{K^2} P(K) + \int_{K_0}^\infty \frac{dK}{K^2} C(K) + \int_{K_0}^F \frac{dK}{K^2} (P(K) - C(K)) \\
&=: \int_0^\infty \frac{dK}{K^2} Q(K) + \int_{K_0}^F \frac{dK}{K^2} (K - F) \\
&\approx \int_0^\infty \frac{dK}{K^2} Q(K) + \frac{1}{K_0^2} \int_{K_0}^F dK\, (K - F) \\
&= \int_0^\infty \frac{dK}{K^2} Q(K) - \frac{1}{K_0^2} \frac{(K_0 - F)^2}{2}.
\end{aligned}$$

One possible discretization of this last expression is

$$VIX^2 = \frac{2}{T} \sum_i \frac{\Delta K_i}{K_i^2} Q_i(K_i) - \frac{1}{T} \left[\frac{F}{K_0} - 1 \right]^2$$

as in the VIX specification (3).

A more accurate formula

- [Fukasawa et al.][7] show that the VIX specification (3) is suboptimal in two respects:
  - After a big move in the underlying, there are typically too few options in the strip to get an accurate estimate to the integral expression for VIX^2.
  - Approximating the integral by a discrete sum leads to errors.

- [Fukasawa et al.][7] consider an alternative approximation using the formula
$$\mathbb{E}[VIX^2] = \int_0^\infty dz\, N'(z)\, \sigma_{BS}^2(z)$$
(with $z = d_-$) that we proved in Lecture 11. They find that this method:
  - gives much more accurate estimates of expected total variance on artificially generated Heston data, and
  - generates better out-of-sample estimates of realized variance.

VIX futures and options

A time-T VIX future is valued at time t as

$$\mathbb{E}_t\left[\sqrt{\mathbb{E}_T\left[\int_T^{T+\Delta} v_s\, ds\right]}\right]$$

where Δ is around one month (or $\Delta \approx 1/12$).

A VIX option expiring at time T with strike K_{VIX} is valued at time t as

$$\mathbb{E}_t\left[\left(\sqrt{\mathbb{E}_T\left[\int_T^{T+\Delta} v_s\, ds\right]} - K_{VIX}\right)^+\right].$$

Remarks on VIX futures and options

- Historically, the options have been much more liquid and actively traded than the futures.

- Both futures and synthetic futures (from put-call parity) used to trade at a substantial premium to the forward-starting variance swap.
  - This arbitrage has come and gone over time.

- Apparently, since the second quarter of 2013, there has been more vega traded in VIX than on SPX, which is now only the second biggest global market for volatility.

VIX futures and options

- Note that we can span the payoff of a forward starting variance swap $\mathbb{E}_t\left[\int_T^{T+\Delta} v_s\, ds\right]$ using VIX options.
- Recall the spanning formula:

$$\mathbb{E}\left[g(S_T)\right] = g(F) + \int_0^F dK\, \tilde{P}(K)\, g''(K) + \int_F^\infty dK\, \tilde{C}(K)\, g''(K).$$

- In this case, $g(x) = x^2$ so

$$\mathbb{E}_t\left[\int_T^{T+\Delta} v_s\, ds\right] = F_{VIX}^2 + 2\int_0^{F_{VIX}} \tilde{P}(K)\, dK + 2\int_{F_{VIX}}^\infty \tilde{C}(K)\, dK.$$

  - F_{VIX} can be computed using put-call parity or observed directly as the VIX futures price.
  - We need to interpolate and extrapolate out-of-the-money option prices to get the *convexity adjustment*.

A real example: 15-Sep-2011

- We compute the fair value of variance swaps from SVI fits to the SPX smiles using the formula

$$\mathbb{E}[VIX^2] = \int_0^\infty dz\, N'(z)\, \sigma_{BS}^2(z).$$

  with $z = d_2$.
- We compare with market variance swap quotes obtained from a friendly investment bank for the same date.

Load some R-code and data

In [1]:
```
download.file(url="http://mfe.baruch.cuny.edu/wp-content/uploads/2015/11/98
unzip(zipfile="9875-12.zip")

library(quantmod)

source("BlackScholes.R")
source("Heston2.R")
source("plotIvols.R")
source("svi.R")
source("sviVarSwap.R")
```

```
Loading required package: xts
Loading required package: zoo

Attaching package: ‘zoo’

The following objects are masked from ‘package:base’:

    as.Date, as.Date.numeric

Loading required package: TTR
Version 0.4-0 included new data defaults. See ?getSymbols.
```

In [3]:
```
# Read in SPX and VIX options data from 15-Sep-2011
load("spxVix20110915.rData")
load("fitQR110915.rData")
```

SPX volatility smiles as of 15-Sep-2011

In [4]:
```
res <- plotIvols(spxOptData,sviMatrix=fitQR )
texp <- res$expiries
```

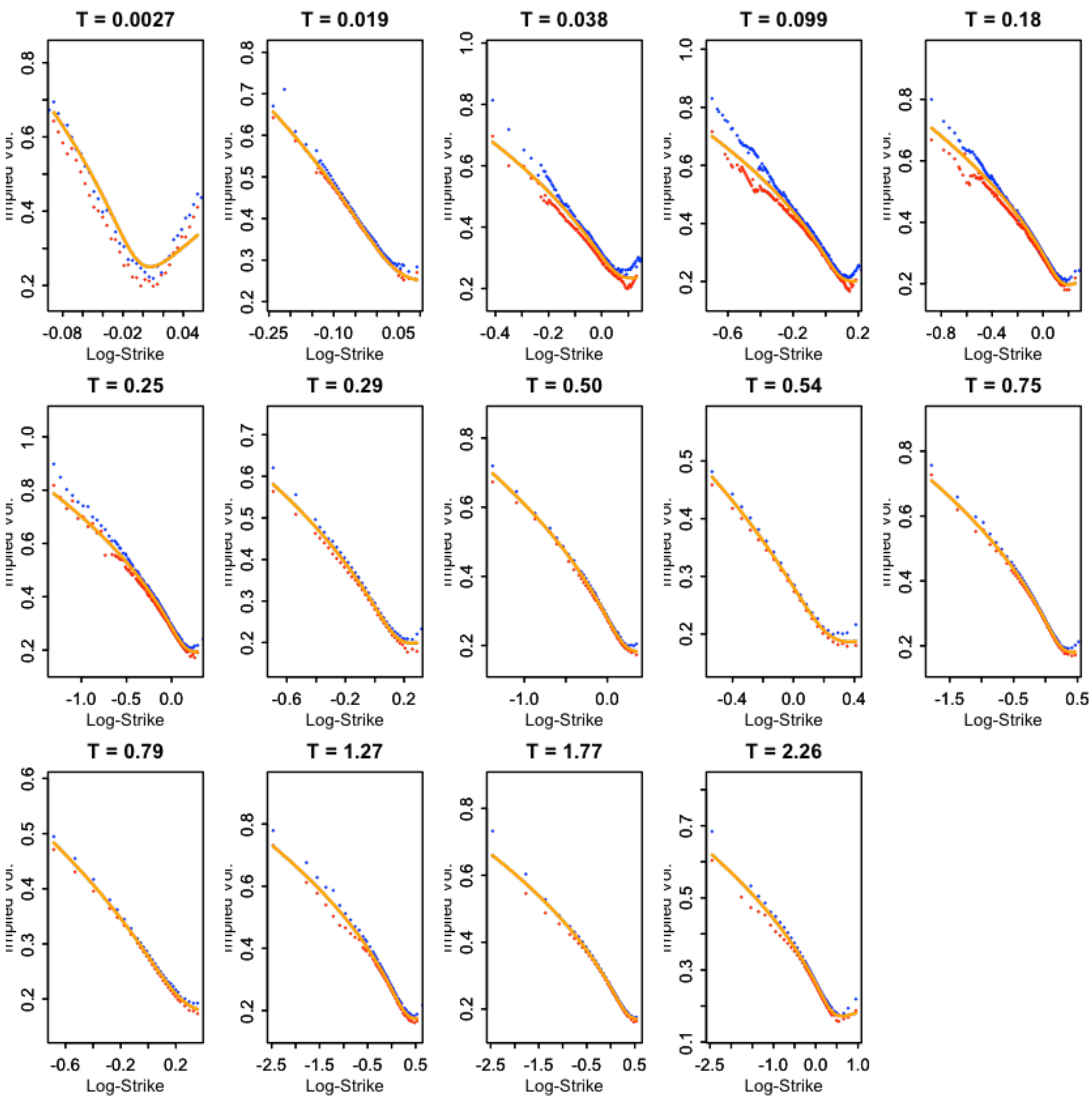


Figure 3: Implied volatility smiles of SPX options as of 15-Sep-2011. SVI fits are in orange.

Variance swaps as of 15-Sep-2011

In [5]:
```
sviVS <- sqrt(sviVarSwap(fitQR,texp=res$expiries))
plot(texp,sviVS,ylab="Variance swap level", xlab="Expiry",type="p",pch=20,c
points(mktVarSwapData$texp,mktVarSwapData$vsBid,col="blue",type="b",lty=2,
points(mktVarSwapData$texp,mktVarSwapData$vsAsk,col="red",type="b",lty=2,pc
```

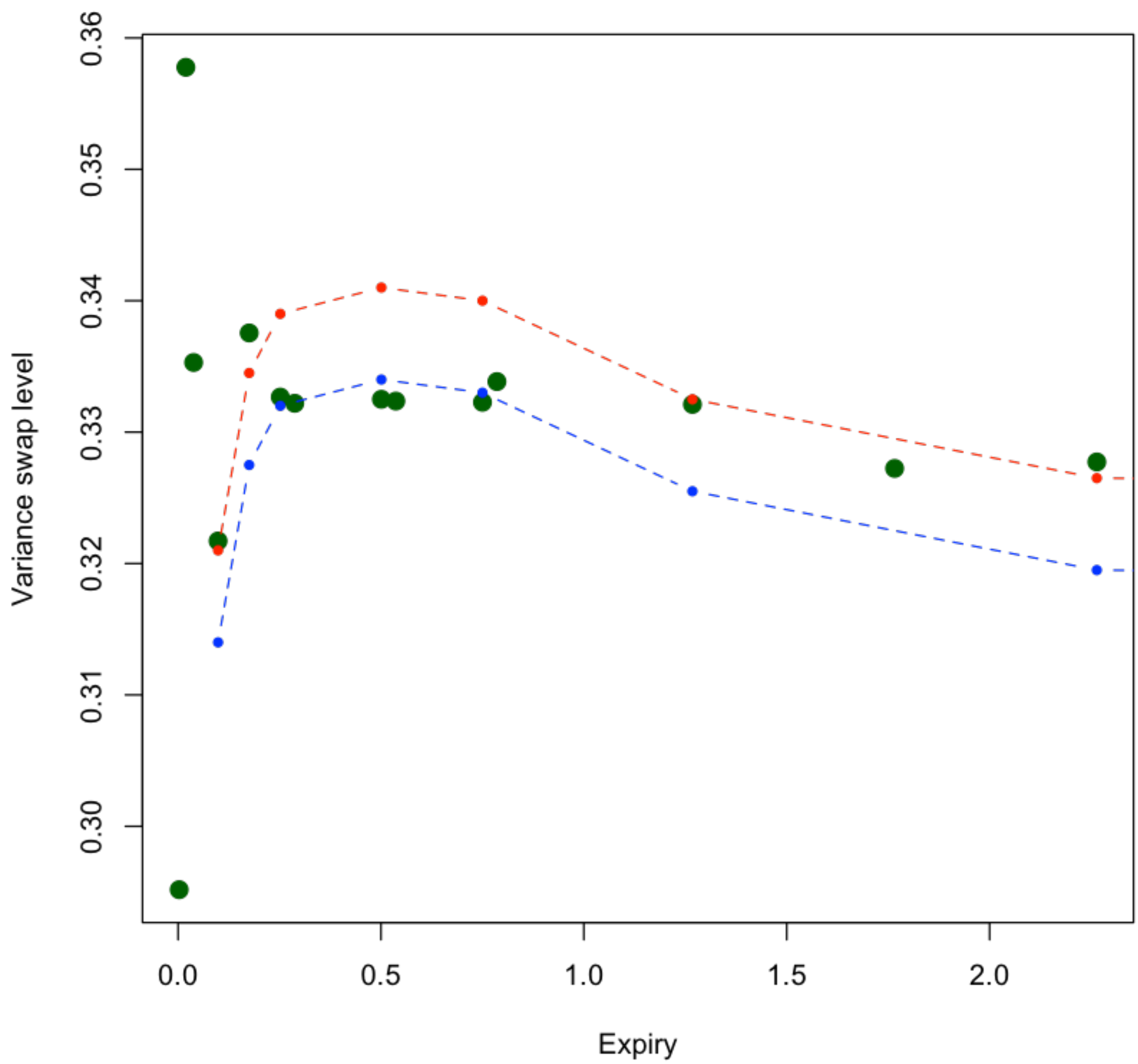


Figure 4: Green dots are computed using the log-strip of SPX options. Blue and red points are bid and ask variance swap quotes from a friendly investment bank.

Forward variance swaps from SPX and VIX options

- Once again, we can compute the fair value of forward starting variance swaps in three ways:
  - Using variance swaps from the SPX log-strip
    - We just computed these variance swaps.

- From the linear strip of VIX options.
- By interpolating market variance swap quotes.

- We now compute the linear VIX strip and compare the three computations as of September 15, 2011.

The VIX volatility smile as of 24-Mar-2010

In [6]:
```
res <- plotIvols(vixOptData)
```

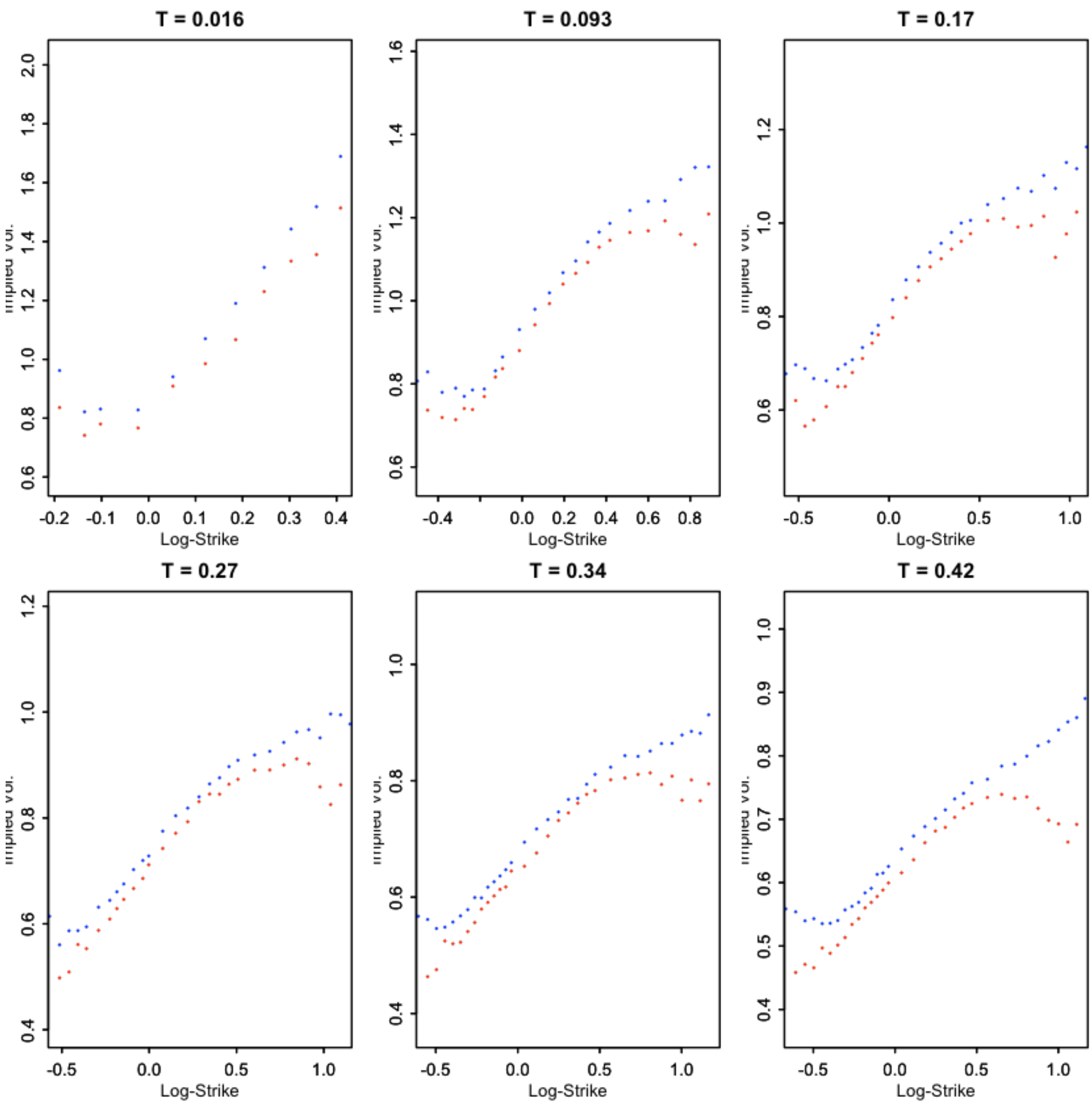


Figure 5: Implied volatility smiles of VIX options as of 15-Sep-2011. Note that all smiles are upward sloping.

Interpolation and extrapolation of VIX smiles

Interpolation and extrapolation of VIX smiles

- We can’t use SVI because VIX smiles are often concave.

- We adopt the Fukasawa robust interpolation/ extrapolation scheme.
  - Monotonic spline interpolation of mid-vols.
  - Extrapolation at constant level.

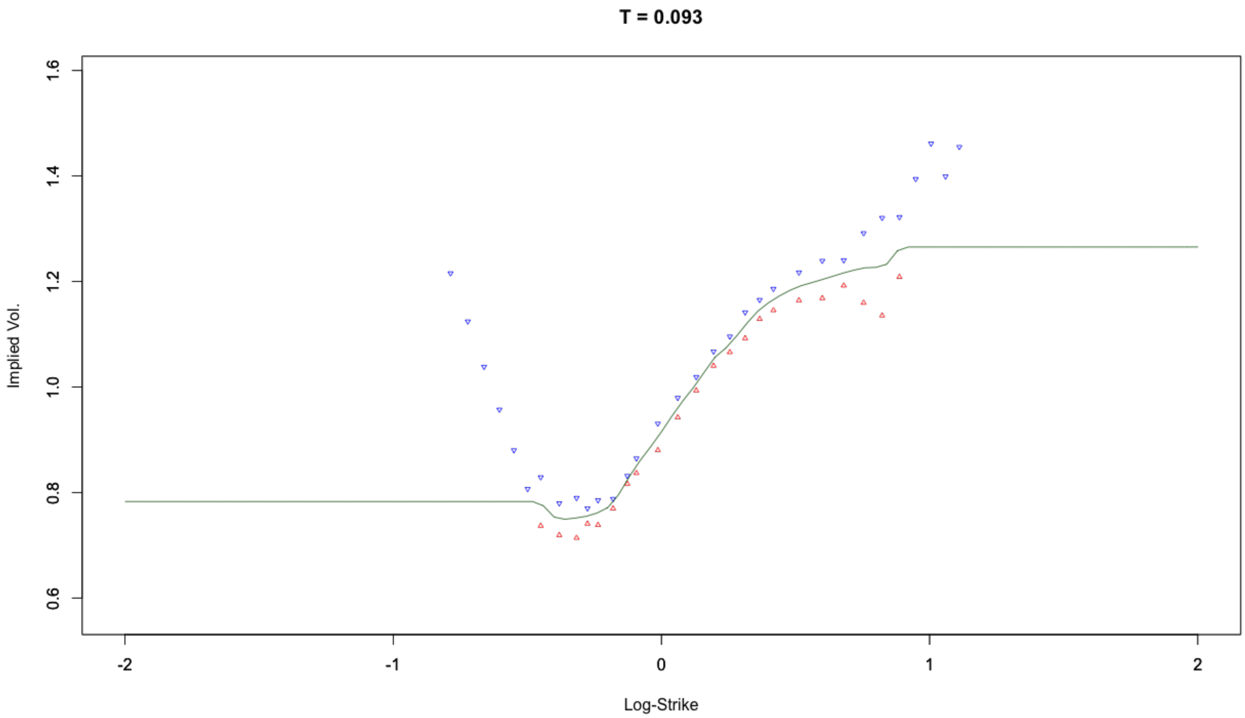


Figure 6: The VIX one month smile with monotonic spline interpolation/ extrapolation.

Forward variance swaps as of 15-Sep-2011

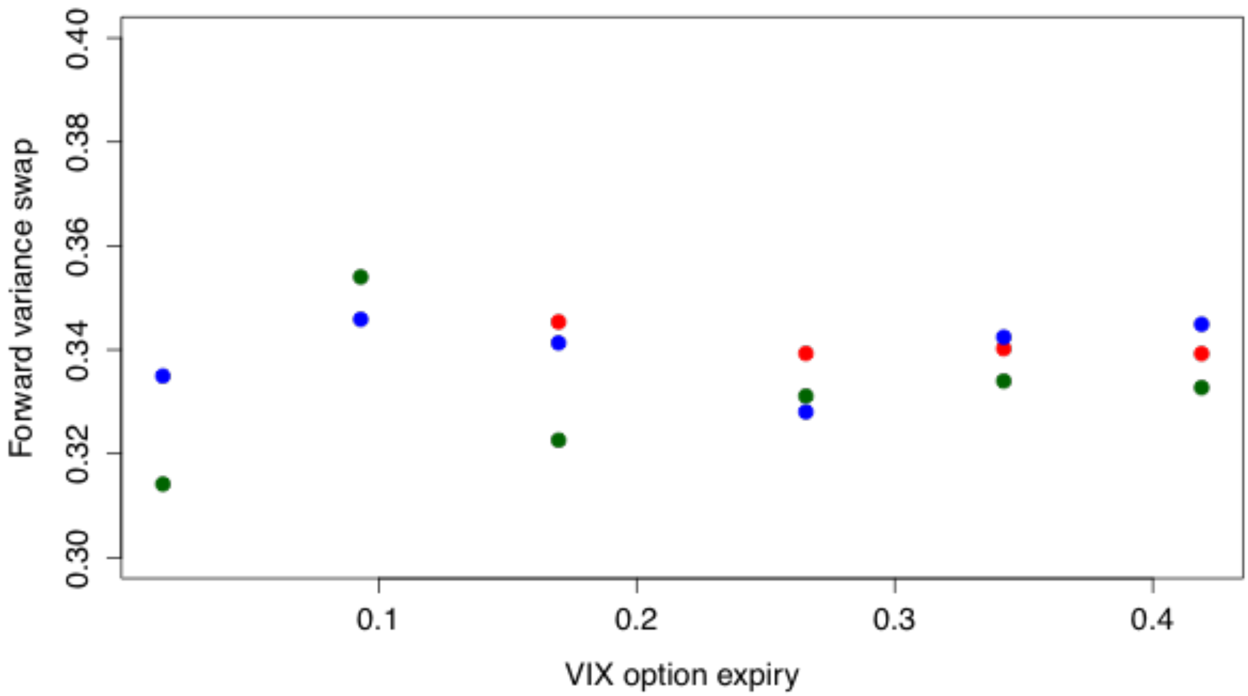


Figure 7: Red dots are forward variance swap estimates from SPX variance swaps; Green dots are interpolation of the SPX log-strip; Blue dots are forward variance swap estimates from the linear VIX option strip.

Consistency of forward variance swap estimates

- Forward variance swap estimates from SPX and VIX are very consistent on this date.
- Sometimes, VIX futures trade at a premium to the forward-starting variance swap.
    - This arbitrage has come and gone over time.
- Taking advantage as a proprietary trader is difficult because you need to cross the bid-ask so often.
    - Buy the long dated variance swap, sell the shorter-dated variance swap.
    - Sell the linear strip of VIX options.
- However, the practical consequence is that buyers of volatility should typically buy variance swaps, sellers should sell VIX.

The VVIX index

- Recall that we may compute the fair value of quadratic variation of $\log S$ from the log-strip of SPX options:

$$\mathbb{E}[\langle \log S \rangle_T] = -2\,\mathbb{E}[\log S_T].$$

- We may also compute the fair value of quadratic variation of $\log VIX$ from the log-strip of VIX options:

$$\mathbb{E}[\langle \log VIX \rangle_T] = -2\,\mathbb{E}[\log VIX_T].$$

- The VVIX index is computed from the log-strip of VIX options in exactly the same way that VIX is computed from SPX options.
  - Roughly speaking, the VVIX index is the fair value of one-month quadratic variation of VIX.

- We then, again roughly speaking, also have a measure of the volatility of volatility.

The fair value of VIX futures

As before, we have $F_{VIX} = \mathbb{E}[VIX]$ and

$$\mathbb{E}[VIX^2] - (\mathbb{E}[VIX])^2 = \mathbb{E}[VIX^2] - F_{VIX}^2 = Var(VIX).$$

Recall that $\mathbb{E}[VIX^2]$ may be computed from a constant strip of VIX options. As for $Var(VIX)$, we have

$$Var(VIX) \approx F_{VIX}^2\, VVIX^2\, T$$

so

$$\mathbb{E}[VIX^2] \approx F_{VIX}^2\, (1 + VVIX^2\, T)$$

which gives yet another way of assessing the fair value of VIX futures.

Why model SPX and VIX jointly

- We had so much success with model-free computations, why should we model SPX and VIX jointly?
  - We may want to value exotic options that are sensitive to the precise dynamics of the underlying such as barrier options, lookbacks or cliquets.
- But is it possible to come up with a parsimonious, realistic model that fits SPX and VIX jointly?
  - And even if we could, would it be possible to calibrate such a model efficiently?

- We will now see that it is indeed possible to do this!

VIX smiles

- As with options on the underlying, knowledge of the following is equivalent:
  - The VIX implied volatility smile
  - VIX option prices
  - The VIX density
  - The VIX characteristic function

- The VIX option smile thus encodes information about the dynamics of volatility in a stochastic volatility model.
  - For example, under stochastic volatility, a very long-dated VIX smile would give us the stable distribution of VIX.

Modeling forward variance

Denote the variance curve as of time t by $\hat{W}_t(T) = \mathbb{E}\left[\int_t^T v_s\, ds \middle| \mathcal{F}_t\right]$. The forward variance $\xi_t(T) := \mathbb{E}\left[v_T \mid \mathcal{F}_t\right]$ is given by

$$\xi_t(T) = \partial_T\, \hat{W}_t(T)$$

A natural way of satisfying the martingale constraint whilst ensuring positivity is to impose lognormal dynamics as in Dupire's (1993) example:

$$\frac{d\xi_t(T)}{\xi_t(T)} = \sigma(T-t)\, dW_t$$

for some volatility function $\sigma(\cdot)$.

We can extend this idea to n-factors.

- This is further extended in our work on rough volatility.

Variance curve models

- The idea (similar to the stochastic implied volatility idea) is to obtain a factor model for forward variance swaps. That is,

  $$\xi_t(T) = G(\mathbf{z}; T-t)$$

  with $\mathbf{z} = \{z_1, z_2, \ldots, z_n\}$ for some factors z_j and some *variance curve functional* $G(\cdot)$.

- Specifically, we want $\mathbf{z}$ to be a diffusion so that

  (4)

  $$d\mathbf{z}_t = \mu(\mathbf{z}_t)\, dt + \sum_j^d \sigma^j(\mathbf{z}_t)\, dW_t^j$$

  - Note that both μ and σ are $n-$dimensional vectors.

Buehler's affine variance curve functional

- Consider the following variance curve functional:

$$G(\mathbf{z};\tau) = z_3 + (z_1 - z_3)\, e^{-\kappa\,\tau} + (z_2 - z_3)\, \frac{\kappa}{\kappa - c}\, (e^{-c\,\tau} - e^{-\kappa\,\tau})$$

  - This looks like the Svensson parametrization of the yield curve.
  - The short end of the curve is given by z_1 and the long end by z_3.
  - The middle level z_2 adds flexibility permitting for example a hump in the curve.

Double mean-reverting (DMR) dynamics

- Buehler's affine variance curve functional is consistent with double mean reverting dynamics of the form:

(5)

$$\begin{aligned} \frac{dS}{S} &= \sqrt{v}\, dW \\ dv &= -\kappa\, (v - v')\, dt + \eta_1\, v^{\alpha}\, dZ_1 \\ dv' &= -c\, (v' - z_3)\, dt + \eta_2\, v'^{\beta}\, dZ_2 \end{aligned}$$

for any choice of $\alpha,\ \beta \in [1/2, 1]$.
  - We will call the case $\alpha = \beta = 1/2$ *Double Heston*,
  - the case $\alpha = \beta = 1$ *Double Lognormal*,
  - and the general case *DMR*[9].

- All such models involve a short term variance level v that reverts to a moving level v' at rate κ. v' reverts to the long-term level z_3 at the slower rate $c < \kappa$.

Discriminating between SV models using VIX options

- Before fitting the DMR model with some arbitrary exponent α, let's compare Double Lognormal and Double Heston fits
  - The idea is that VIX option prices should allow us to discriminate between models.

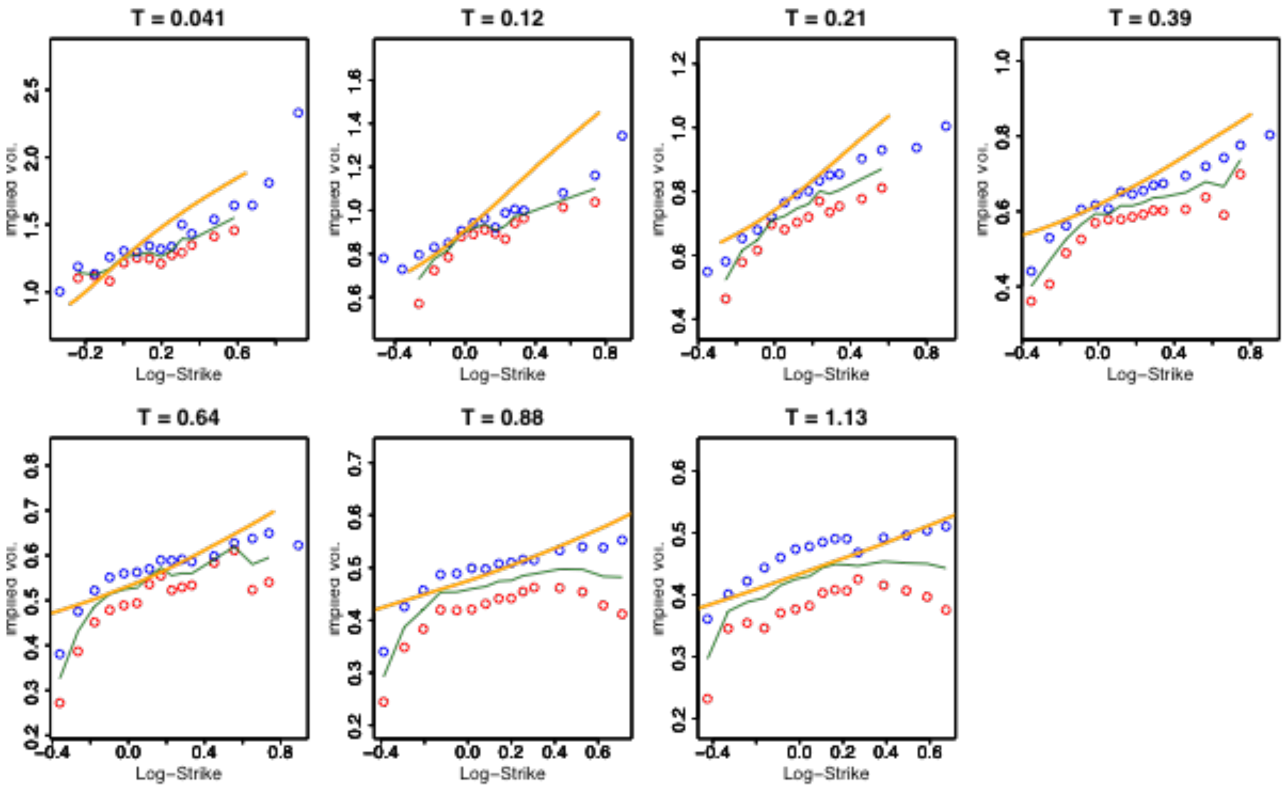
- Specifically, we want to assess:
  - Fit quality
  - Parameter stability

Fit of Double Lognormal model to VIX options

As of 03-Apr-2007, from Monte Carlo simulation with parameters

$$z_1 = 0.0137;\ z_2 = 0.0208;\ z_3 = 0.0421;\ \kappa = 12;\ \xi_1 = 7;\ c = 0.34;\ \xi_2 = 0.94;$$

we get the following fits (orange lines):

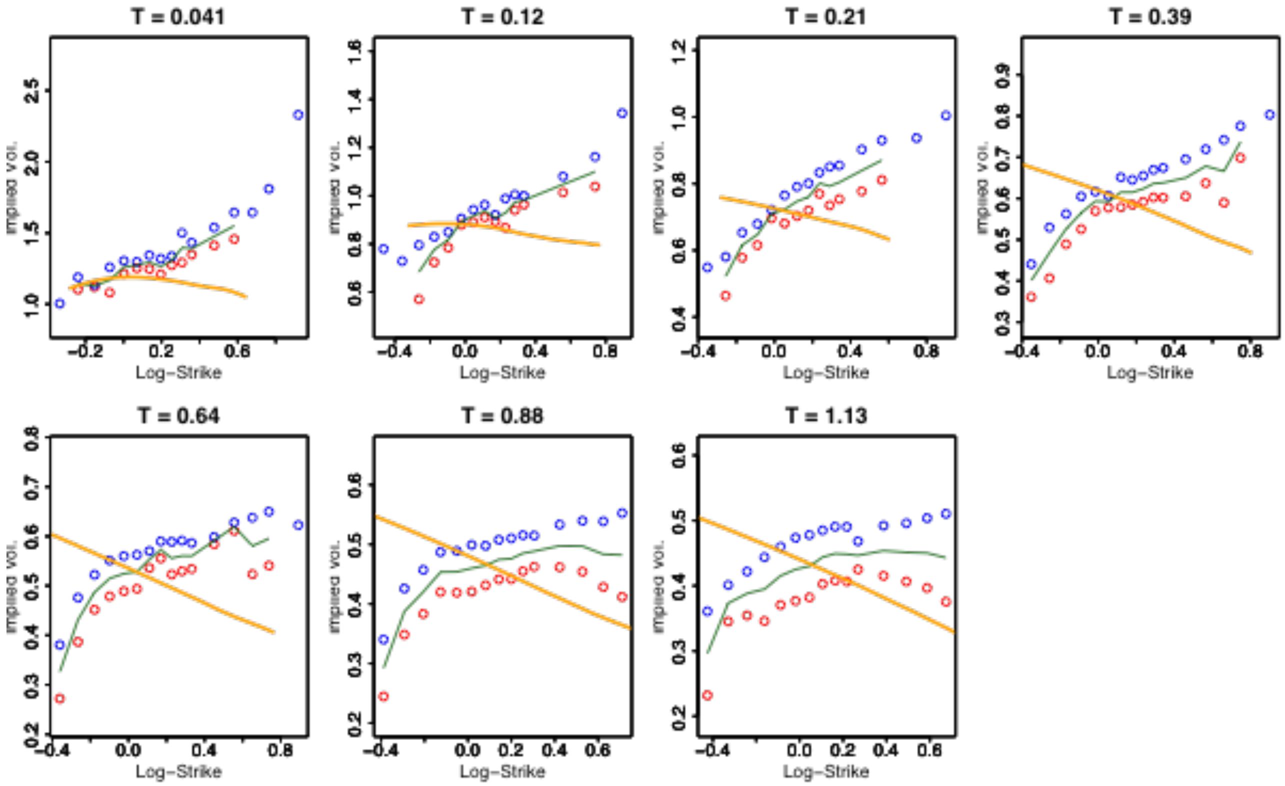


Fit of Double Heston model to VIX options

As of 03-Apr-2007, from Monte Carlo simulation with parameters

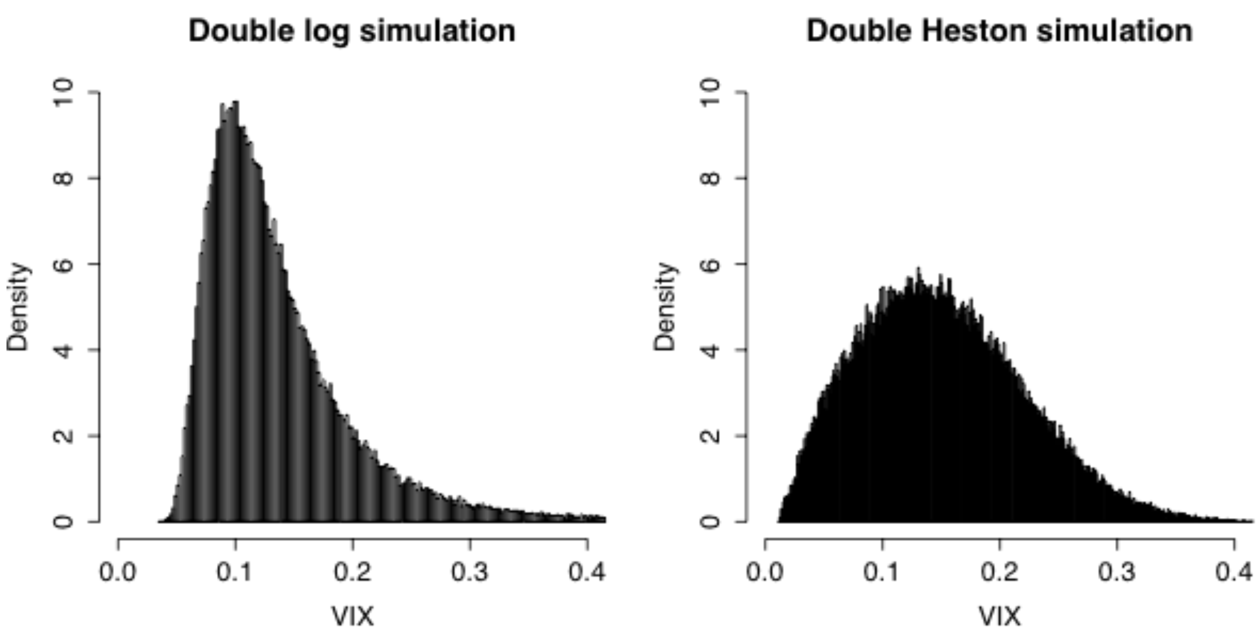
$$z_1 = 0.0137;\ z_2 = 0.0208;\ z_3 = 0.0421;\ \kappa = 12;\ \xi_1 = 0.7;\ c = 0.34;\ \xi_2 = 0.14;$$

we get the following fits (orange lines):



In terms of densities of VIX

- When we draw the densities of VIX for the last expiration ($T = 1.13$) under each of the two modeling assumptions, we see what's happening:



- In the (double) Heston model, v_t spends too much time in the neighborhood of $v = 0$

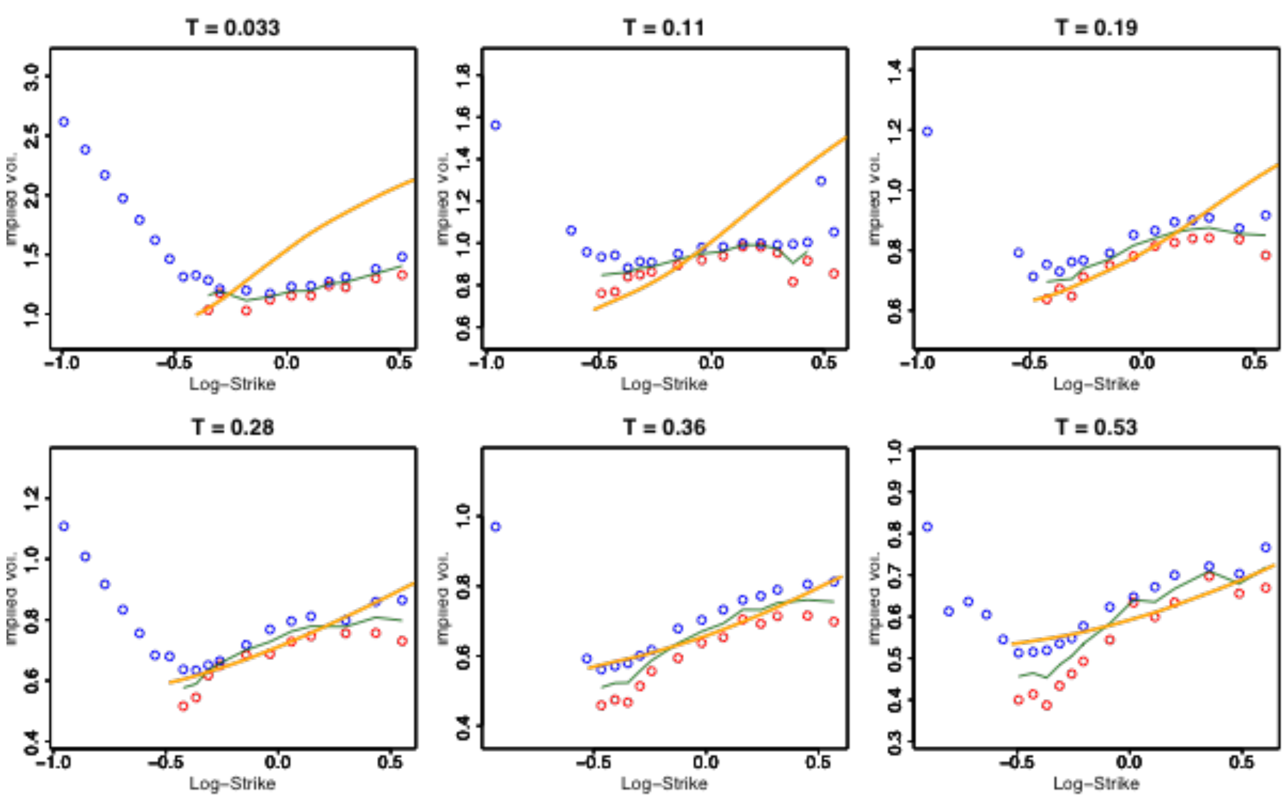
and too little time at high volatilities.

Parameter stability

- Suppose we keep all the parameters unchanged from our 03-Apr-2007 fit. How do model prices compare with market prices at some later date?
- Recall the parameters:
  - Lognormal parameters:
    $$\kappa = 12;\ \xi_1 = 7;\ c = 0.34;\ \xi_2 = 0.94;$$
  - Heston parameters:
    $$\kappa = 12;\ \xi_1 = 0.7;\ c = 0.34;\ \xi_2 = 0.14;$$
- Specifically, consider 09-Nov-2007 when volatilities were much higher than April.
  - We re-use the parameters from our April fit, changing only the state variables z_1 and z_2.

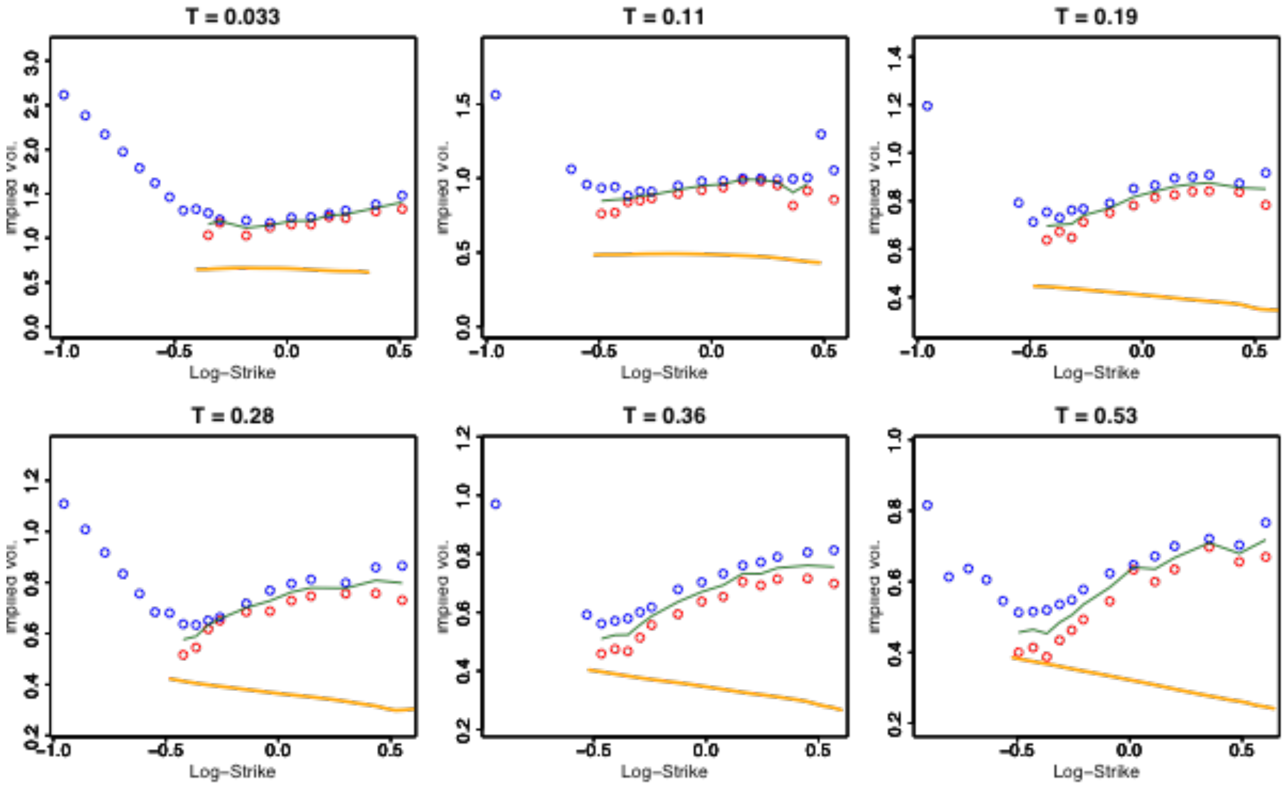
Double Lognormal fit to VIX options as of 09-Nov-2007

With $z_1 = 0.0745$, $z_2 = 0.0819$ we get the following plots (model prices in orange):



Double Heston fit to VIX options as of 09-Nov-2007

With $z_1 = 0.0745$, $z_2 = 0.0819$ we get the following plots (model prices in orange):



Observations

- The Double Lognormal model clearly fits the market better than Double Heston.
  - Not only does Double Lognormal fit better on a given day, but parameters are more stable over time.
  - The fair value of a put on VIX struck at 5% should be negligible but Double Heston says not.
  - VIX option prices are inconsistent with Double Heston dynamics.

- This is consistent with our stylized fact that volatility is lognormal.

- The VIX smile remains the same shape...

Computations in the DMR model

- One drawback of the DMR model is that calibration is not easy
  - No closed-form solution for European options exists so finite difference or Monte Carlo methods need to be used to price options.
  - Calibration using conventional techniques is therefore slow.

- In [Bayer, Gatheral and Karlsmark][2], the DMR model is calibrated using the Ninomiya-Victoir Monte Carlo scheme.
  - Joint calibration of the model to SPX and VIX options is possible in less than 5

seconds.

Estimation of κ_1, κ_2, θ and ρ_{23}

- In the DMR model, the fair strike of a variance swap is given by the expression

  (6)

  $$\begin{aligned}
  &\mathbb{E}\left[\int_t^T v_s\, ds \middle| \mathcal{F}_t\right] \\
  &= \theta\,\tau + (v_t - \theta)\frac{1 - e^{-\kappa_1\,\tau}}{\kappa_1} \\
  &\quad + (v_t' - \theta)\frac{\kappa_1}{\kappa_1 - \kappa_2}\left\{\frac{1 - e^{-\kappa_2\,\tau}}{\kappa_2} - \frac{1 - e^{-\kappa_1\,\tau}}{\kappa_1}\right\}
  \end{aligned}$$

  which is affine in the state variables v_t and v_t'.

- Fixing θ, κ_1 and κ_2, and given daily variance swap estimates, time series of v_t and v_t' may be imputed by linear regression.
  - Optimal values of θ, κ_1 and κ_2 are obtained by minimizing mean squared differences between the fitted and actual variance swap curves.

Daily model fitting

- The model parameters κ_1, κ_2, θ and ρ_{23} are considered fixed. They are obtained from historical variance swap data.
- The state variables v_t and v_t' are obtained by linear regression against the fair values of variance swaps proxied by the log-strip.
  - Arbitrage-free interpolation and extrapolation of the volatility surface is achieved using the SVI parameterization in [Gatheral and Jacquier][10].

- The volatility-of-volatility parameters ξ_1 and ξ_2 are obtained by calibrating the DMR model to the market prices of VIX options (using NVs).
- The correlation parameters ρ_{12} and ρ_{13} are then calibrated to SPX options.

Term structures of VIX volatility and skew

In [7]:

```
# Figure 8. VIX ATM volatility and skew
par(mfrow=c(1,2))
plot(res$expiries,res$atmVol,col="red",ylab="Implied vol.",xlab="Expiry")
curve(res$atmVol[3]*(res$expiries[3]/x)^.25,from=0,to=.5,add=T,col="blue")
plot(res$expiries,res$atmSkew, col="red",ylab="Vol. skew", xlab="Expiry")
curve(res$atmSkew[3]*(res$expiries[3]/x)^.72,from=0,to=.5,add=T,col="dark g
par(mfrow=c(1,1))
```

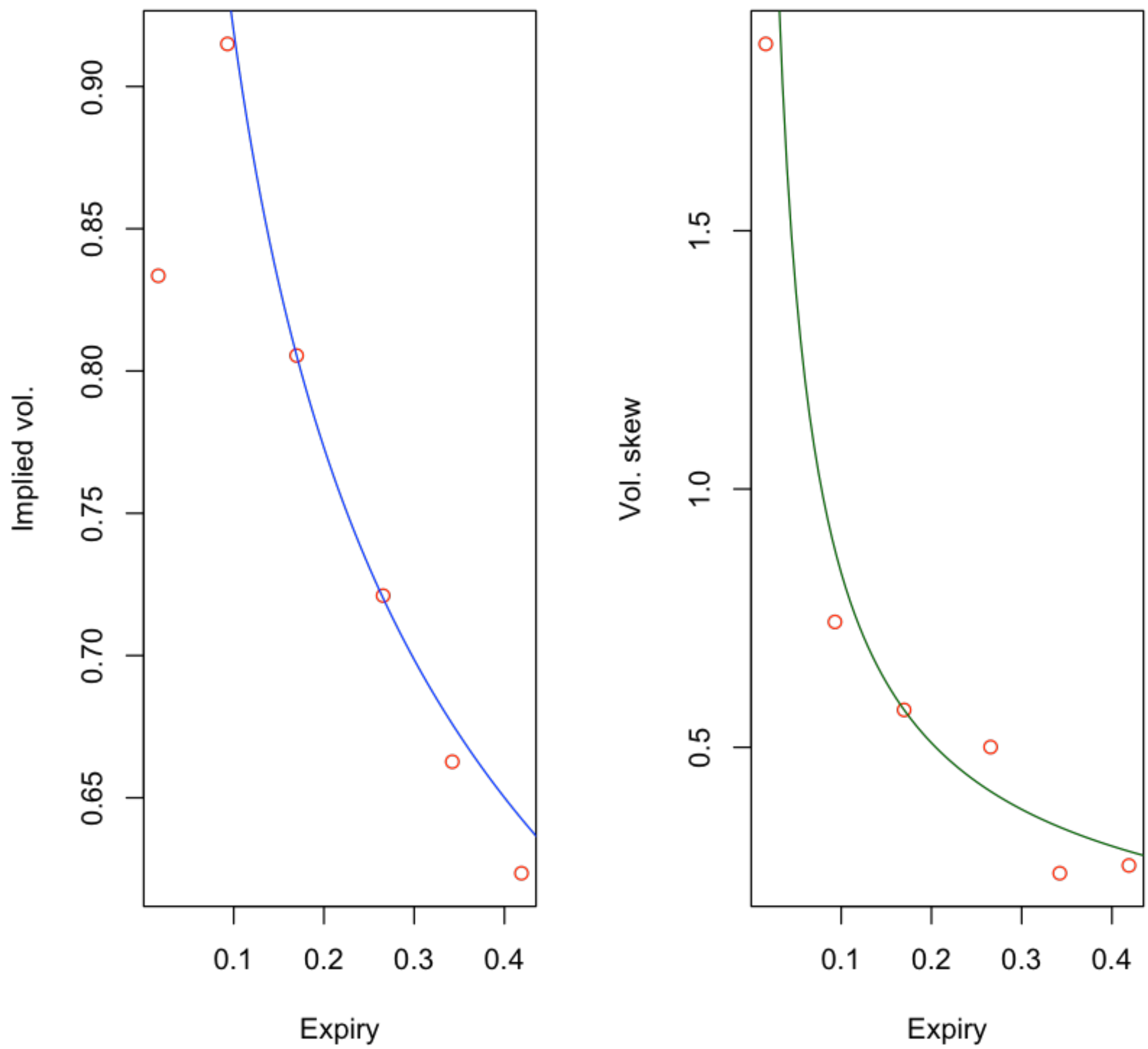


Figure 8: Term structures of VIX volatility and skew as of 15-Sep-2011. The blue line is the power law $1/t^{1/4}$ and the dark green line the power-law $1/t^{.72}$.

Fitting SPX options

Recall the DMR dynamics:

$$\left.\begin{aligned} \frac{dS}{S} &= \sqrt{v}\,dW \\ dv &= -\kappa\,(v - v')\,dt + \eta_1\,v^{\alpha}\,dZ_1 \\ dv' &= -c\,(v' - z_3)\,dt + \eta_2\,v'^{\beta}\,dZ_2 \end{aligned}\right|$$

with $\mathbb{E}[dZ_1\,dZ_2] = \rho\,dt$, $\mathbb{E}[dZ_1\,dW] = \rho_1\,dt$, $\mathbb{E}[dZ_2\,dW] = \rho_2\,dt$.

- Fitting the historical variance swap curves gives us c, κ and z_3.
- Fitting the VIX options prices gives us the parameters η_1, η_2, α, β and ρ.
- Using Monte Carlo, we fit to SPX options prices to get ρ_1 and ρ_2.

VIX fit as of September 15, 2011

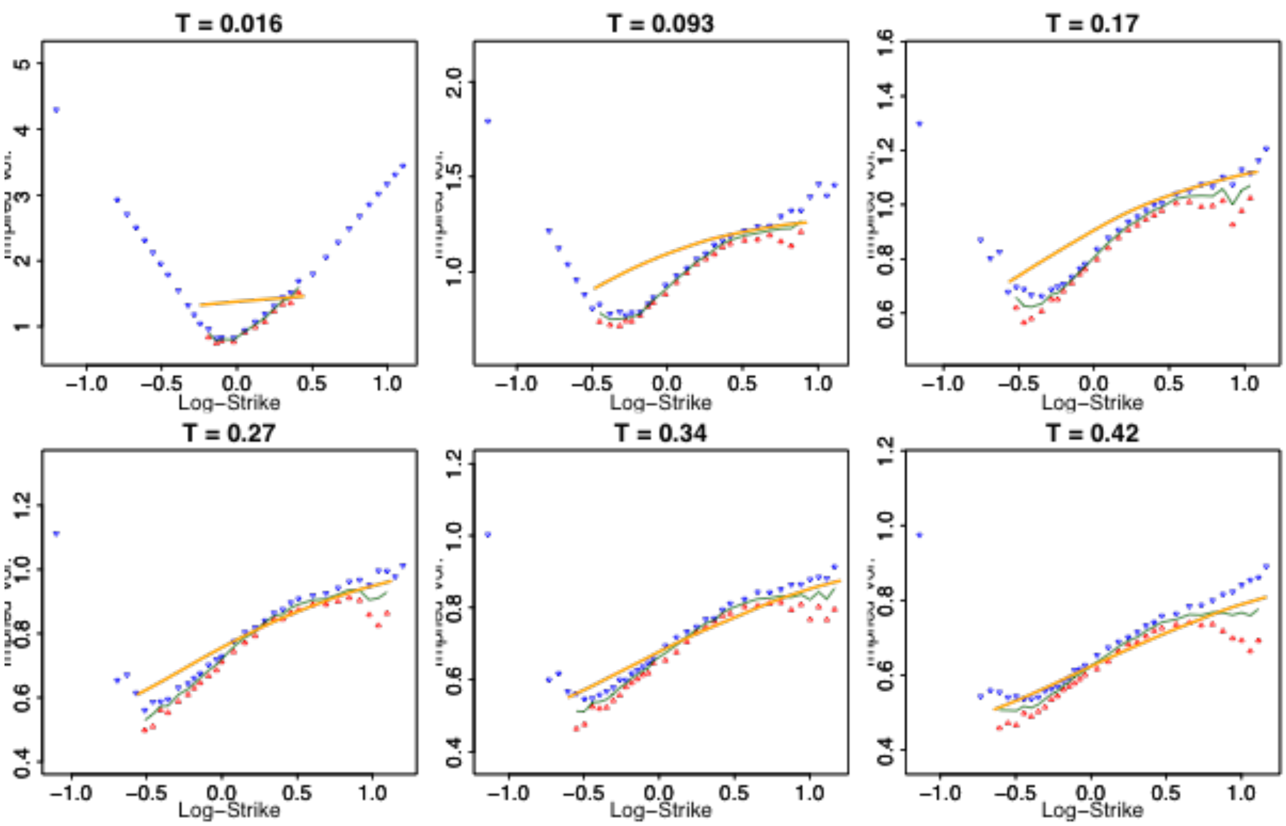
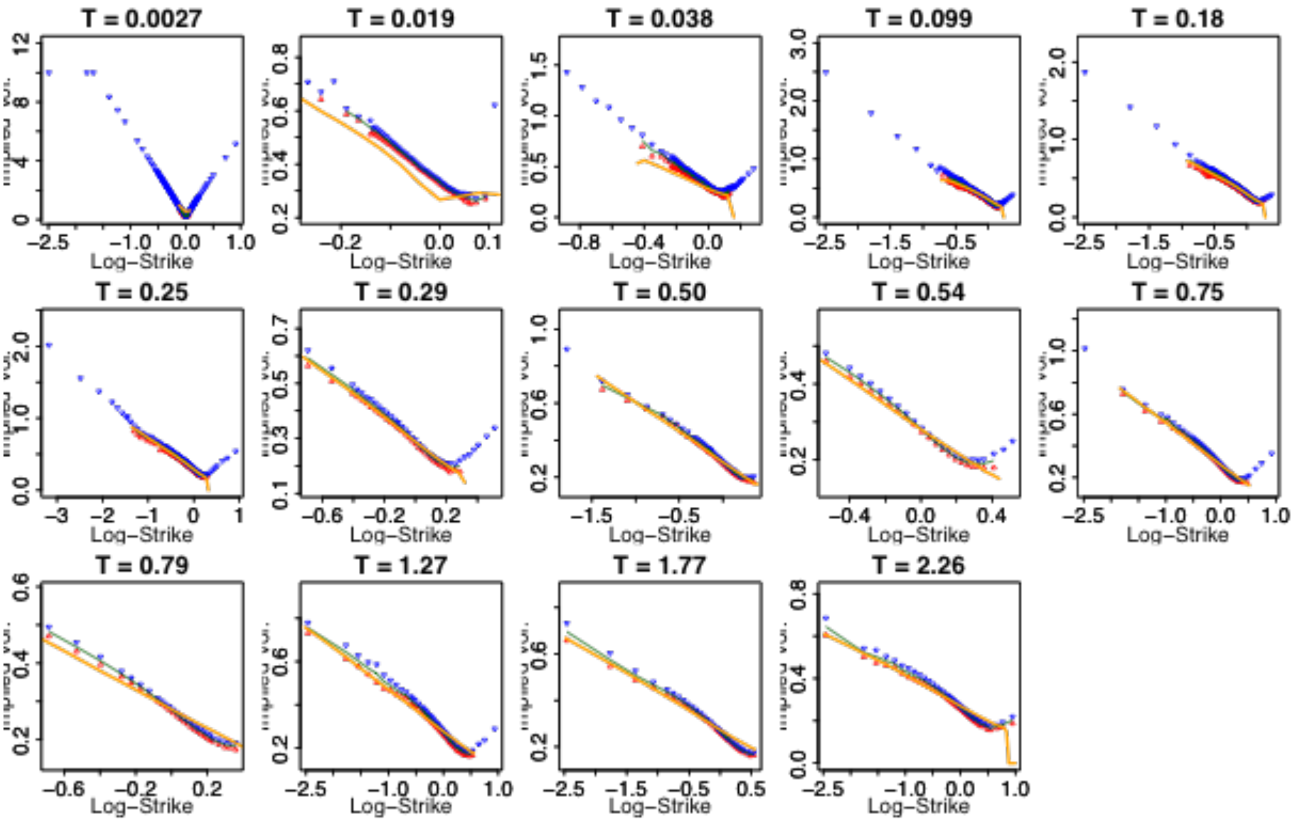


Figure 9: VIX smiles as of September 15, 2011: Bid vols in red, ask vols in blue, and model fits in orange.

SPX fit as of September 15, 2011



Parameters

- Collating all our calibration results:

Parameter	Value
κ	5.50
c	0.10
z_3	0.078
ξ_1	2.6
α	0.94
ξ_2	0.45
β	0.94
ρ	0.59
ρ_1	-0.90
ρ_2	-0.70

"More tractable" models

- To compute option prices, and so for calibration, models such as DMR require Monte Carlo simulation.

- We now explore two tractable alternatives:
  - Sepp (2008)
  - Cont and Kokholm (2013)

The Sepp model

The model of [Sepp][11] assumes the following dynamics:

$$\begin{aligned} \frac{dS_t}{S_t} &= \sigma(t)\,\sqrt{v_t}\,dZ_t \\ dv_t &= -\kappa\,(v_t - 1)\,dt + \epsilon\,\sqrt{v_t}\,dW_t + J\,dN_t. \end{aligned}$$

where N is a Poisson process with intensity γ. Jumps are exponentially distributed with density

$$\frac{1}{\eta}e^{-J/\eta}.$$

- Roughly speaking, Heston with a positive jump in the volatility. $\sigma(t)$ is a piecewise constant function calibrated to match the term structure of VIX futures.

Parameter estimates from July 25, 2007 are:

Parameter	Value
κ	2.26
ϵ	1.66
η	2.54
γ	0.31

- We see that the mean jump size is huge with significant probability!

- Sepp computes the characteristic function with respect to $x = \log S$, v, and the integrated variance $I = \int_t^T \sigma^2(t)\,v_s\,ds$. (That's cool!)

- Using this characteristic function, he can compute the prices of VIX options and SPX options using (at most) 2-dimensional numerical integration.

The Cont-Kokholm model

Forward starting variance swaps are modeled directly. Let

$$V_t^i = \int_{T_i}^{T_{i+1}} \xi_t(u)\, du$$

where $\xi_t(u)$ is the forward variance curve. Then in the model of [Cont and Kokholm][5] ,

$$V_t^i = V_0^i \exp\left\{ \int_0^t \mu_s^i\, ds + \omega \int_0^t e^{-\kappa_1 (T_i - s)}\, dW_s + \sum_{j=0}^{N_t} e^{-\kappa_2 (T_i - \tau_j)}\, Y_j \right\}$$

where N is a Poisson process with intensity λ, the τ_j are its jump times and the Y_j are jumps with distribution F independent of W.

- Jumps Y_j could for example be lognormally or exponentially distributed.
  - with simultaneous jumps in SPX of roughly $b_i x$ where x is the size of the jump in volatility.
- There is an accurate approximation for VIX.
  - VIX options are priced in closed form.
- The positive skew in VIX options comes from positive jumps in volatility.
- The negative skew in SPX comes from simultaneous negative jumps.

- Calibrated parameters correspond to roughly
  - 3 or 4 jumps per year
  - a $+50\%$ mean jump in variance
  - a -10% mean jump in the index
- VIX options give the size and frequency of jumps
- SPX options are used to fit parameters that
  - relate the index jump to the volatility jump
  - determine a piecewise constant diffusion coefficient

VIX-based ETNs

- There exist various ETNs that contain positions in Treasuries and VIX futures:
  - VXX: Long short-term VIX futures (Barclays)
  - VXZ: Long medium-term VIX futures (Barclays)
  - XIV: Inverse VIX Short Term ETN (Credit Suisse)
  - ZIV: Inverse VIX Medium Term ETNs (Credit Suisse)
  - VIIX: Long VIX Short Term ETNs (Credit Suisse)
  - VIIZ: Long VIX Medium Term ETNs (Credit Suisse)
  - TVIX: $2\times$ Long VIX Short Term ETNs (Credit Suisse)

- TVIZ: $2\times$ Long VIX Medium Term ETNs (Credit Suisse)

VXX

From the iPath website:

> The S&P 500 VIX Short-Term Futures™ Index Total Return offers exposure to a daily rolling long position in the first and second month VIX futures contracts. On each business day, a fraction of the first month VIX futures holding is sold and an equal notional amount of the second month future is bought, maintaining a constant weighted average maturity of one month.

VXZ

From the iPath website:

> The S&P 500 VIX Mid-Term Futures™ Index Total Return offers exposure to a daily rolling long position in the fourth, fifth, sixth and seventh month VIX futures contracts. On each business day, a fraction of the fourth month VIX futures holding is sold and an equal notional amount of the seventh month future is bought, maintaining a constant weighted average maturity of five months.

ETFs according to Avellaneda and Zhang

Following [Avellaneda and Zhang][1], denote the spot price of the underlying by S_t, the price of the leveraged ETF (LETF) by L_t, and leverage ratio by β.

Assume that

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

Then assuming zero interest rates, dividends and borrowing costs, we have

$$\frac{dL_t}{L_t} = \beta\,\frac{dS_t}{S_t}$$

Leveraged ETF valuation

Applying Itô’s Lemma,

$$d\log S_t = \frac{dS_t}{S_t} - \frac{1}{2}\,\sigma_t^2\,dt$$

and

$$d\log L_t = \frac{dL_t}{L_t} - \frac{1}{2}\,\beta^2\,\sigma_t^2\,dt = \beta\,\frac{dS_t}{S_t} - \frac{1}{2}\,\beta^2\,\sigma_t^2\,dt$$

Then

(7)

$$d\log L_t - \beta\,d\log S_t = -\frac{1}{2}\,(\beta^2 - \beta)\,\sigma_t^2\,dt$$

Integrating (7) gives

(8)

$$\frac{L_t}{L_0} = \left(\frac{S_t}{S_0}\right)^{\beta}\,\exp\left\{\frac{1}{2}\,(\beta - \beta^2)\,\int_0^t\,\sigma_s^2\,ds\right\}$$

Volatility exposure of leveraged ETF

- We see from (7) and (8) that leveraged ETFs are short volatility.
- For example:
  - If $\beta = 2$,

$$d\log L_t - \beta\,d\log S_t = -\sigma_t^2\,dt$$

  - Likewise, if $\beta = -1$,

$$d\log L_t - \beta\,d\log S_t = -\sigma_t^2\,dt$$

- This is because the LETF has to buy the underlying when it goes up and sell it when it goes down.

Rebalancing

- Wlog, suppose the starting value of the LETF is $N_0 = \$1$.
- At the close of day 1, the notional is

$$V_1 = 1 + \beta\,\frac{\Delta S_1}{S_0}$$

- After rebalancing, at the close of day 2, the value is

(9)

$$V_2 = N_1\,\left(1 + \beta\,\frac{\Delta S_2}{S_1}\right) = \left(1 + \beta\,\frac{\Delta S_2}{S_1}\right)\,\left(1 + \beta\,\frac{\Delta S_1}{S_0}\right)$$

- On the other hand if there is no rebalancing the value would be

(10)

$$V_2' = \left(1 + \beta\,\frac{\Delta S_1 + \Delta S_2}{S_0}\right)$$

dollars worth of underlying.

- Subtracting (10) from (9), we obtain

$$\begin{aligned} V_2 - V_2' &= \beta \left\{ \frac{\Delta S_2}{S_1} - \frac{\Delta S_2}{S_0} + \beta \, \frac{\Delta S_2}{S_1} \, \frac{\Delta S_1}{S_0} \right\} \\ &= \beta \, (\beta - 1) \, \frac{\Delta S_1}{S_0} \, \frac{\Delta S_2}{S_1} \end{aligned}$$

from which we conclude that the rebalancing at the end of day 1 involved buying

$$\beta \, (\beta - 1) \, \frac{\Delta S_1}{S_0}$$

dollars worth of the underlying.

- A leveraged ETF is therefore short gamma: it buys(sells) the underlying if the price rises(falls).

Regression of VXX vs VIX

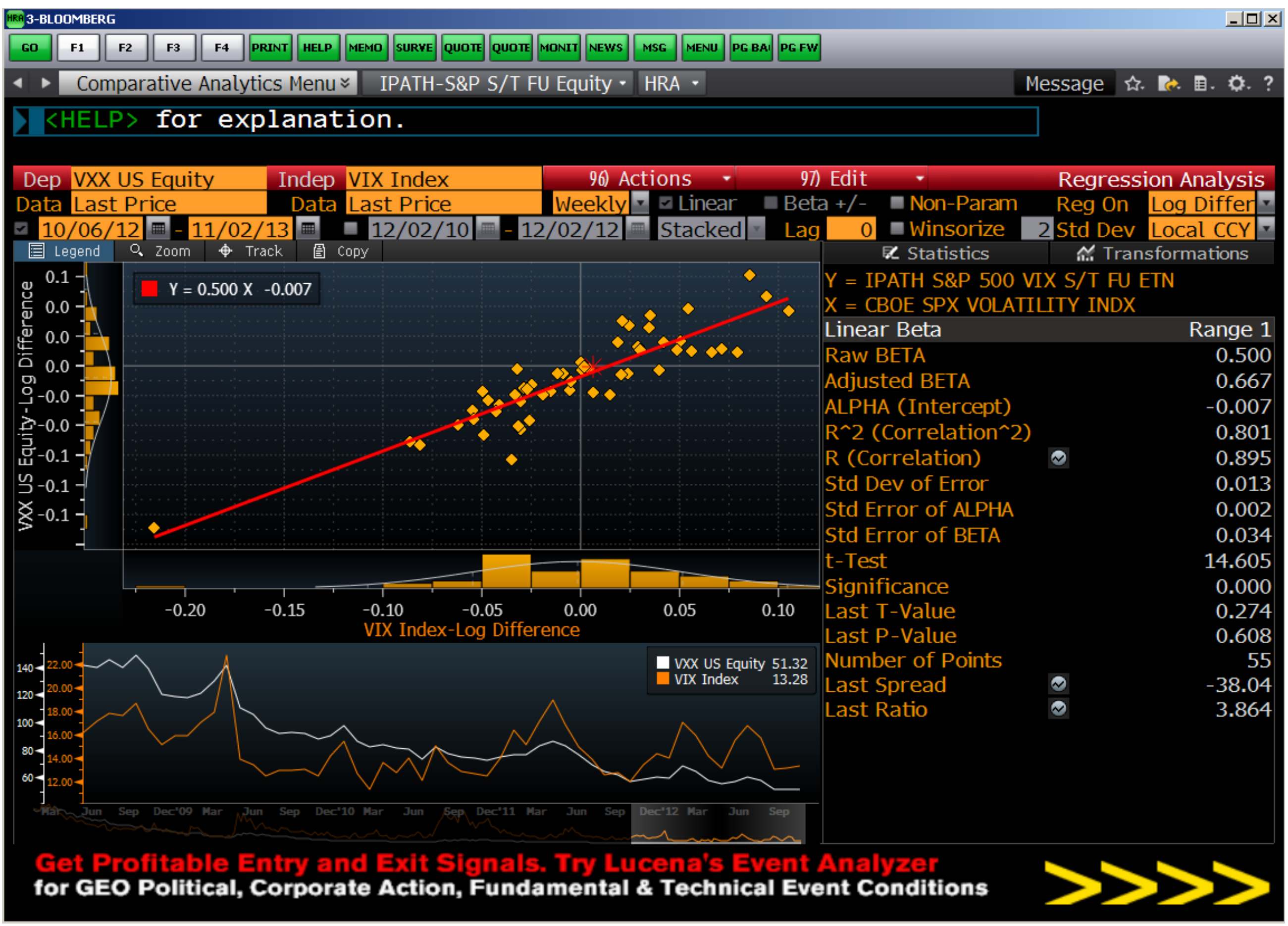


Figure 10: Regression of VXX vs VIX

Regression of XIV vs VXX

Figure 11: Regression of XIV vs VXX

XIV vs VXX

In [8]:
```
getSymbols(c("^VIX","VXX","VXZ","XIV"),from="1927-01-01")

mm <- specifyModel(Cl(XIV)~Cl(VXX))
xivData <-modelData(mm)  #quantmod function automatically aligns data from

# Repeat over period with no adjustments
xiv1 <- xivData$Cl.XIV["2013-11-08/2014-11-30"]
vxx1 <- xivData$Cl.VXX["2013-11-08/2014-11-30"]
plot(xiv1,main=NA)
lines(vxx1,col="red")
```

```
    As of 0.4-0, 'getSymbols' uses env=parent.frame() and
 auto.assign=TRUE by default.

 This  behavior  will be  phased out in 0.5-0  when the call  will
 default to use auto.assign=FALSE. getOption("getSymbols.env") and
```

```
 getOptions("getSymbols.auto.assign") are now checked for alternate de
faults

 This message is shown once per session and may be disabled by setting
 options("getSymbols.warning4.0"=FALSE). See ?getSymbols for more deta
ils.
```

'VIX' 'VXX' 'VXZ' 'XIV'

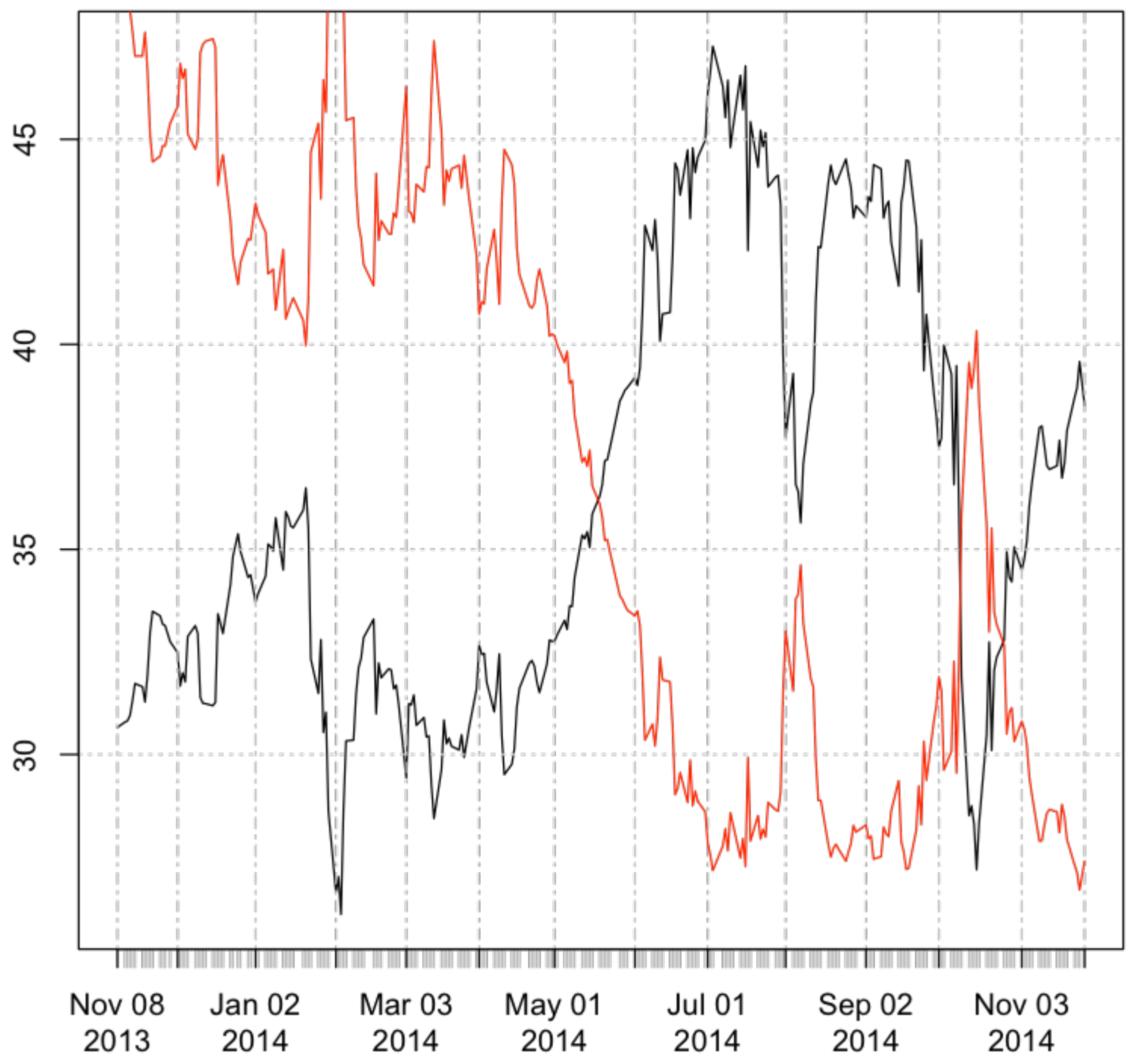


Figure 12: VXX in red; XIV in black

XIV vs VXX

In [9]:
```
plot(vxx1/as.numeric(vxx1[1])+xiv1/as.numeric(xiv1[1]) ,main=NA)
```

Figure 13: One dollar invested in each of VXX and XIV

Summary

- Volatility derivatives are now very actively traded, not least because from a practical perspective, valuation is well-understood.
- We saw that although the relationship between the valuation of options on variance and European option prices is currently only partially understood, it's pretty clear from a practical perspective how to value them.
- On the theoretical side, connections between the valuation of volatility derivatives and

the dynamics of the underlying process continue to be investigated.

References

1. ^ M Avellaneda and S Zhang, Path-dependence of Leveraged ETF returns, *SIAM J. Financial Math.* **1** 586-603 (2010).
2. ^ Christian Bayer, Jim Gatheral, and Morten Karlsmark, Fast Ninomiya-Victoir calibration of the double-mean-reverting model, *Quantitative Finance* **13**(11) 1813-1829 (2013).
3. ^ Mathias Beiglböck, Peter Friz and Stephan Sturm, Is the minimum value of an option on variance generated by local volatility?, *SIAM J. Financial Math.* **2** 213–220 (2010).
4. ^ Peter Carr and Roger Lee, Hedging variance options on continuous semimartingales, *Finance and Stochastics* **14** 179–207 (2010).
5. ^ Rama Cont and Thomas Kokholm, A consistent pricing model for index options and volatility derivatives, *Mathematical Finance* **23**(2) 248-274 (2013).
6. ^ Bruno Dupire, Volatility derivatives modeling, available at http://www.math.nyu.edu/~carrp/mfseminar/bruno.ppt (http://www.math.nyu.edu/~carrp/mfseminar/bruno.ppt) (2005).
7. ^ M Fukasawa, I Ishida, N Maghrebi, K Oya, M Ubukata, and K Yamazaki, Model-Free Implied Volatility: From Surface to Index, *International Journal of Theoretical and Applied Finance* **14**(4) 433–463(2010).
8. ^ Jim Gatheral, *The Volatility Surface: A Practitioner's Guide*, John Wiley and Sons, Hoboken, NJ (2006).
9. ^ Jim Gatheral, Consistent modeling of SPX and VIX options, *Bachelier Congress*, London (2008).
10. ^ Jim Gatheral and Antoine Jacquier, Arbitrage-free SVI volatility surfaces, *Quantitative Finance* **14**(1) 59-71 (2014).
11. ^ Artur Sepp, VIX option pricing in a jump-diffusion model, *Risk Magazine* 84-89 (April 2008).


```
In [ ]:
```