University of Illinois, Urbana-Champaign
FIN 513 - Spring 2018

Professor Tim Johnson
tcj@illinois.edu

Lecture Note 7.3: Model-Free Implied Volatility

This note tells the story of one of the most useful and creative achievement in financial engineering in the last 25 years.

As we have discussed many times, there is a fatal flaw in Black-Scholes implied volatility: it has no consistent interpretation in terms of the *true* volatility of the returns of the underlying asset unless that volatility is constant – which it never is.

Yet for investors, risk manager, and regulators it would be extremely useful to be able to use the information in option prices to gauge current uncertainty. However, as we have seen, different models of uncertainty (i.e. different specifications of stochatic volatility) will lead to different options prices. Therefore the same option prices (from the market) coupled with different models will imply different levels of current volatility and expected uncertainty.

The beautiful discovery is that there is actually a way to measure expected uncertainty for options prices that *does not depend on the stochastic assumptions!*

Outline:

I. The Log Contract
II. VIX
III. Example
IV. Extensions

I. The Log Contract.

- The story begins with an observation due to Anthony Neuberger in 1994.

- He asked the following question:

  ▶ Suppose there is an asset whose current price is S_0 and whose time-T forward price is $F_{0,T}$ and we are short a derivative contract that will require us to pay $\log(S_T/F_{0,T})$ at time T.

  ▶ Further suppose we choose to engage in a dynamic hedging strategy between now and T in which every Δt (e.g. every day) we hold $1/F_{t,T}$ forwards.

  ▶ *How much money will we make?*

- This seems like a funny question. But it is pretty easy to answer. Let's just do the accounting.

- Here are our trades.

  Day 0: Buy $\frac{1}{F_{0,T}}$ forwards at price $F_{0,T}$.

  Day 1: Buy $\frac{1}{F_{1,T}} - \frac{1}{F_{0,T}}$ forwards at price $F_{1,T}$.

  ⋮ ⋮

  Day $\mathbf{T-1}$: Buy $\frac{1}{F_{T-1,T}} - \frac{1}{F_{T-2,T}}$ forwards at price $F_{T-1,T}$.

- When we get to time T all of these trades settle. Our profits from each trade are:

  Day 0: $[\frac{1}{F_{0,T}}]\ \ (F_{T,T} - F_{0,T}) +$

  Day 1: $[\frac{1}{F_{1,T}} - \frac{1}{F_{0,T}}]\ \ (F_{T,T} - F_{1,T}) +$

  $\vdots \qquad\qquad\qquad \vdots$

  Day $\mathbf{T-1}$: $[\frac{1}{F_{T-1,T}} - \frac{1}{F_{T-2,T}}]\ \ (F_{T,T} - F_{T-1,T})$.

- Please verify that adding all these terms up gives

$$\sum_{t=0}^{T-1} \left(\frac{F_{t+1,T}}{F_{t,T}} - 1 \right)$$

  or

$$\sum_{t=0}^{T-1} \left(\frac{\Delta F_{t+1,T}}{F_{t,T}} \right).$$

- Now, what is the difference between the quantity in parentheses, $\Delta F/F$, and $\Delta \log F$?
  - ▶ If we let Δt get very small so that we are talking about infinitessimals, then, for any diffusion process, we have from Ito's lemma:

$$d \log(F) = \frac{1}{F}\ dF - \frac{1}{2}\sigma^2\ dt$$

    where σ is the volatility of F.

- ▶ So, for small time intervals, we can replace $\Delta F/F$ by $\Delta \log F + \frac{1}{2}\sigma^2\ \Delta t$.
- ▶ Notice that this is the *model free* step: I did not assume anything about the σ here.
  - ∗ It could be stochastic – obeying any specification at all.
  - ∗ It could even have jumps.

- When we make the substitution for $\Delta F/F$ in our hedging profit expression, our total P&L is

$$\sum_{t=0}^{T-1} \left(\Delta \log F_{t+1,T} + \frac{1}{2}\sigma_t^2\ \Delta t \right)$$

which reduces to

$$\log(F_{T,T}) - \log(F_{0,T}) + \frac{1}{2}\sum_{t=0}^{T-1} \sigma_t^2\ \Delta t$$

since all the intermediate log terms cancel out.

- Now the first terms here are the same as $\log(F_{T,T}/F_{0,T}) = \log(S_T/F_{0,T})$. And this is exactly the payout that we are short.
- So the total payoff at T from our position is just the second term:

$$\frac{1}{2}\sum_{t=0}^{T-1} \sigma_t^2\ \Delta t \qquad \text{or} \qquad \frac{1}{2}\int_0^T \sigma_t^2\, dt.$$

- Notice also that the $1/F$ hedging strategy that we engaged in was self-financing. In fact, it was costless.

- Thus, a short position in the log contract plus a costless strategy generates a payoff that is equal to (half) the realized variance between now and T.

- **Neuberger's conclusion:** The market value of the log contract must be equal to the minus one half times the market value of a claim to the realized total variance.

- Or, the value of the log contract must be $-T/2$ times the market value of a claim to the realized average variance.

- Now let us define the model free implied variance as:

$$MFIV^2 = \frac{1}{T}\,\mathrm{E}_0^Q\left[\int_0^T \sigma_t^2\,dt\right]$$

  where $\mathrm{E}^Q[\,]$ denotes expectation taken with respect to the risk-neutral measure.

  - The present value of that expectation is also equal to the price of contract that pays the realized average variance at T.

- If L_0 denotes the price of the log contract, we have thus deduced

$$L_0 = -\frac{T}{2}\,B_{0,T}\;MFIV^2.$$

- So, if we could get market participants to tell us the value of the log contract, we would have measured (up to a multiplicative constant) the quantity we were after.

II. VIX

- The most important volatility that everyone wants to measure is that of "the market portfolio".
  - ▶ For a well-diversified investor, in fact, this quantity is the ONLY risk that really matters.

- Since the early 1990s, the CBOE had published daily summary statistics about Black-Scholes implied volatility of options on the S&P 500.

- But they faced a problem: *which options?*
  - ▶ As we've discussed, every strike and maturity would give them a different answer.
  - ▶ They used the most liquid ones, which are short-dated and at-the-money. But those don't capture the beliefs expressed in out-of-the-money puts.
  - ▶ But forming some kind of weighted average also seems kind of arbitrary

- So they were very interested when they heard about Neuberger's idea about using the value of the (theoretical) log contract to summarize the market's beliefs.

- Then came another problem: *How can it be implemented?*

- The log contracts are not actually traded and no dealer quotes prices in them.

- There was no reason to think if they were listed that any institutions would actually be interested in them.
- If we can't observe prices for the log contract, we've just replaced one measurement problem with another.
- But Neuberger (together with Mark Britten-Jones) had an answer.
  - Butterflies!
- We learned early on in the course that if we have prices for a lot of ordinary European calls (or puts) we can infer the prices of butterflies, and then construct arbitrary payoff functions from them.
  - Well, the CBOE does have the for data on calls and puts on the S&P500.
  - So they should be able to synthesize something very close to a log contract's payoff by an appropriate portfolio of butterflies.
  - Or, since the butterflies are themselves just made up of calls and puts, we should be able to construct a weighting scheme that just sums over all of them.

- That's what they did. And thus the modern VIX was born.

- Let's try to understand the weighting scheme for synthesizing the log contract.

- Consider the idealized case with infinite strike prices available:
  - If $b(k)$ denotes the price of a butterfly centered at strike price k, then we learned that the value of the log contract must be
  $$L_0 = \sum_k \log(k/F)\, b(k).$$
  (where I'm writing F for $F_{0,T}$.)

  - Now expand the butterflies into their component options.

$$\sum_k \log(k/F)\, b(k) = \sum_k \log(k/F)\, [p(k-\Delta k)-2p(k)+p(k+\Delta k)]/\Delta k.$$

  Recall the butterflies are the same whether we use puts or calls to build them.

  - Actually, it will be more convenient to use both calls and puts, but on different ranges. So write the sum:

$$\sum_{k=\Delta k}^{F_{0,T}} \log(k/F)\ \frac{p(k-\Delta k) - 2p(k) + p(k+\Delta k)}{\Delta k}$$

$$+ \sum_{k=F_{0,T}}^{\infty} \log(k/F)\ \frac{c(k-\Delta k) - 2c(k) + c(k+\Delta k)}{\Delta k}$$

- Now take the first term. Re-write this sum by grouping together terms involving the same put.
- (I'll supress the Δk and F and just index everything by integers.)

$$\begin{array}{rllllll}
\log(1) & (p(0) & -2\,p(1) & +\,p(2)) & & & \\
+\log(2) & & (p(1) & -2\,p(2) & +\,p(3)) & & \\
+\log(3) & & & (p(2) & -2\,p(3) & +\cdots & \\
+\vdots & & & \vdots & & &
\end{array}$$

$$= p(1)[\log(0){-}2\log(1){+}\log(2)] + p(2)[\log(1){-}2\log(2){+}\log(3)] + \cdots$$

  * For example, if $\Delta k = 0.01$ then the terms involving puts with $k = 100.00$ are

    $$p(100.00) * [\log(99.99) - 2\log(100.00) + \log(100.01)]$$

  * Also, to be careful near zero, the formula is assuming $\log(1)p(0) - \log(0)p(1)$ goes to zero.

    - Mathematically, sufficient conditions for this are that $p(x)d\log(x)/dx \to 0$ and $\log(x)dp/dx \to 0$.

- This will be fine as long as the left tail of the risk neutral distribution goes to zero exponentially, for example.

- Now, bringing back the Δk that I suppressed, our first summation is

$$\sum_{k=\Delta k}^{F_{0,T}} p(k) \; \frac{[\log(k-\Delta k) - 2\log(k) + \log(k+\Delta k)]}{\Delta k^2} \; \Delta k$$

  * Inside each of the logs there was a $(1/F_{0,T})$, but we can drop them because each term in the summation has

$$-\log(F_{0,T}) + 2\log(F_{0,T}) - \log(F_{0,T}) = 0$$

- Likewise, I can re-arrange the call sum as

$$\sum_{k=F_{0,T}}^{\infty} c(k) \; \frac{[\log(k-\Delta k) - 2\log(k) + \log(k+\Delta k)]}{\Delta k^2} \; \Delta k$$

- As the strikes become infinitely close together, you will recognize that the fractional terms are just going to become second derivatives.

- And we know from calculus that $\frac{d^2 \log(x)}{dx^2} = -\frac{1}{x^2}$.

- So we can conclude that our double sum is equal to

$$-\left(\int_0^{F_{0,T}} \frac{p(k)}{k^2} \; dk + \int_{F_{0,T}}^{\infty} \frac{c(k)}{k^2} \; dk \right).$$

- So far we have deduced that this integral gives us the value, L_0, of the log contract
- Since we deduced above that $L_0 = -\frac{T}{2}\, B_{0,T}\, MFIV^2$, we conclude

$$MFIV^2 = \frac{2}{T}\, B_{0,T}^{-1} \left(\int_0^{F_{0,T}} \frac{p(k)}{k^2}\, dk + \int_{F_{0,T}}^{\infty} \frac{c(k)}{k^2}\, dk \right).$$

- That equation is in fact what the CBOE's measure tries to approximate using finite sums instead of integrals over an infinite range.
  - ▶ The squre root of that <u>is</u> VIX.
- You can compute this for any maturity option. And the available maturities change every day.
  - ▶ The primary VIX is supposed to always represent options with 30 days to maturity. Since these don't usually exist, the CBOE just linearly interpolates the (squared) VIXes computed from two expirations on either side of 30 days.
  - ▶ They also do this for 9-day, 3-month, and 6-month horizons. These are called (in order) VXST, VXV, and VXMT.
- As a practical matter, they are limited by only having finitely many strikes.

- The first integral is the numerically sensitive part because it involves both numerator and denominator going to zero as k goes to zero.
  - The exchange has listed some puts with extremely low strike prices to help in this direction.

- Exact details of the biases due to (a) upper and lower limits not being 0 and infinity and (b) finite Δk have been analyzed by Jiang and Tian (2007) and (2009).
  - They also suggested some numerical corrections.
  - As far as I know, CBOE has not adopted them yet.

- Still it is remarkable how much you can do with butterflies.
  - It is fascinating that a contract whose payoff depends on the realized, random path of volatility can be replicated by a *static* portfolio of puts and calls.

III. Example Computations.

- We have seen that VIX (squared) measures model-free risk-neutral expected integrated variance for the S&P 500 for a fixed time horizon.

- You might be asking yourself some questions about this quantity, such as:

(A) How does the average variance relate to the *current* instantaneous variance? (And how do their dynamics relate to each other?)

(B) How does the risk-neutral expected variance relate to the *true* expected variance?

- I thought it would be helpful to do an explicit computation (NOT model-free) to address these.

- Let's assume the true instantaneous variance $v_t = \sigma_t^2$ obeys

$$dv = \kappa(\bar{v} - v)dt + svdW^v$$

This is a little different from the σ specification we used before, but it's the same idea.

- Now I will just state a handy fact that can be proved with a little stochastic analysis. For this process,

$$\mathrm{E}[v_u|v_t] = v_t \, e^{-\kappa(u-t)} + \bar{v}(1 - e^{-\kappa(u-t)}).$$

- ▶ This just says that your forecast for volatility at a time horizon of $(u-t)$ in the future is a linearly weighted average of today's value and the long-run mean.

- ▶ And the weights decline exponentially at rate given by κ.

- So for this model it is easy to compute the average expectation (which is VIX):

$$\mathrm{E}\left[\frac{1}{\tau}\int_t^{t+\tau} v_u\ du|v_t\right] = \frac{1}{\tau}\int_t^{t+\tau} \mathrm{E}\left[v_u|v_t\right]\ du$$

$$= \frac{1}{\tau}\int_t^{t+\tau} [v_t\ e^{-\kappa(u-t)}\ +\ \bar{v}(1-e^{-\kappa(u-t)})]\ du$$

$$= \bar{v}\ +\ \frac{1}{\tau}[(v_t-\bar{v})\int_t^{t+\tau} e^{-\kappa(u-t)}\ du = \bar{v}\ +\ (v_t-\bar{v})\frac{1-e^{-\kappa\tau}}{\kappa\tau}$$

$$= w(\tau)v_t + (1-w(\tau))\bar{v}$$

where $w = w(\tau) = (1-e^{-\kappa\tau})/\kappa\tau$.

- ▶ Like the fixed-horizon expectation, the average expectation is a weighted average of today's value and the long-run value.

- ▶ If we call this weighted expectation $V_t(\tau)$, then we can ask how it changes over time while the horizon τ stays fixed – as it does for VIX.

- ▶ This is simple because $dV_t = w(\tau)dv_t$

- To compute VIX, however, we were supposed to use risk-neutral expectations.

- Well, we've learned how to change the process to back and forth between the two measures:

$$
\begin{aligned}
dv &= [\kappa(\bar{v} - v) - \lambda\, s\, v]\; dt \;+\; s\, v\, dW^v \\
&= (\kappa + \lambda\, s)\left(\frac{\kappa\bar{v}}{\kappa + \lambda\, s} - v\right)\; dt \;+\; s\, v\, dW^v \\
&= \kappa^\star(v^\star - v)\; dt \;+\; s\, v\, dW^v.
\end{aligned}
$$

- Very interesting! This tells us that to "risk-neutralize" this model we have to adjust the long-run level of variance <u>and</u> the speed of mean reversion to the new values that I denoted $\kappa^\star$ and $v^\star$.
  - ▶ Recall that under the Vasicek specification we only had the second adjustment.

  Q: If the market price of variance risk is negative, which way do these adjustments go?

  Q: Are they big?

  Q: What happens if $\kappa + \lambda\, s \leq 0$?

- Returning to our earlier calculation, now the VIX index will be a weighted average of v_t and $v^\star$ with weight, $w^\star$, given by

$$
w^\star = \frac{1 - e^{-\kappa^\star \tau}}{\kappa^\star \tau} \qquad \text{instead of} \qquad \frac{1 - e^{-\kappa\tau}}{\kappa\tau}.
$$

- And the same substitution go into $dV = w^\star dv$ – the (true) dynamics of the risk-neutral expected average variance.
  - And, since $VIX = \sqrt{V} = \sqrt{w^\star}\sqrt{v}$, the dynamics of VIX are the same as those of true volatility multiplied by a constant.
  - In particular, we can conclude that VIX should be perfectly correlated with changes in true volatility.

- The specific calculations here will be different, of course, if you start with a different model of the true variance dynamics.
- As an exercise, you might want to re-do the steps above with a different assumption.
  - For example, let $Z = \log(\sigma) = 0.5\ \log(v)$ and suppose
$$dZ = \kappa(\bar{Z} - Z)\,dt + s\,dW^Z$$
    i.e. Z follows an Ornstein-Uhlenbeck process.
- You will still find the same result from my calculation that *the risk-neutral expectation is biased upward* relative to the true expectation.
- And the bias is proportional to the market price of volatility risk.
- Notice that this bias does <u>not</u> imply that options markets are inefficient or irrational in any sense.

- In fact, there is a lot of on-going research that tries to measure the gap between the best true volatility forecast for the market and the VIX.
  - This is termed the "volatility risk premium".
  - It seems that this quantity is itself time varying.
  - And its variations may capture important information about crash fears in the economy.

IV. Extensions.

- Model-free VIX has been so successful in gaining widespread acceptance as the correct measure of risk-neutral expected volatility that the CBOE has even introduced cash-settled futures on VIX and options on those futures.
  - ▶ Using our earlier results, a good exercise would be to derive expressions for the no-arbitrage futures (or forward) price of VIX.

  These futures are the basis of very popular VIX-based ETFs and inverse ETFs.

- They have also begun to compute model-free implied volatilities on other indexes and on commodities like gold and crude oil.

- Every major stock exchange in the world now computes a VIX clone for its own set of indexes.

- CBOE now even computes VIX on VIX !
  - ▶ See `http://www.cboe.com/micro/VVIX/`

- Most recently, they have borrowed another idea from Neuberger and started to compute model-free <u>skewness</u>.
  - ▶ This is a measure of the asymmetric (risk-neutral) crash risk in the market.
  - ▶ See `http://www.cboe.com/micro/SKEW/`

V. Summary

- It is a very cool fact that we can derive the risk-neutral expected cumulative variance from butterfly prices without ever having to specify what we think the *true* dynamics of volatility are.

- Having done so, we still have to be careful in interpreting what VIX means. It is not the instantaneous volatility of the market. And its dynamics are not the same as the dynamics of σ_t.

- Most important, it is not an unbiased forecast of future market volatility: it embeds a very substantial negative risk premium.
  - ▶ The discovery of the market volatility risk premium is one of the most important recent developments in financial research.