Why is VIX a fear gauge?

Peter Carr
NYU Tandon School of Engineering, 12 MetroTech Brooklyn, NY 11201, USA
E-mail: pc73@nyu.edu

Abstract. VIX is a widely followed volatility index constructed from the market prices of out-of-the-money (OTM) puts and calls written on the S&P500. VIX is often referred to as a fear gauge. While the market prices of OTM puts clearly reflect the fear that the S&P500 will drop, about half of the options used in constructing VIX are actually OTM calls. The market prices of these OTM calls clearly reflect greed rather than fear. In this note, we offer several explanations as to why VIX can properly be regarded as a fear gauge.

Keywords: VIX, volatility, fear gauge

1. Introduction

VIX is a widely followed index constructed from the market prices of out-of-the-money (OTM) puts and calls written on the S&P500. The V in VIX stands for volatility. It is well known that since its redesign in 2003, VIX approximates the cost of synthesizing a newly issued 30 day variance swap under a set of widely accepted assumptions.

Besides its well known role as a volatility index, VIX is often referred to as a fear gauge. While the market prices of OTM puts clearly reflect the fear that the S&P500 will drop, about half of the options used in constructing VIX are actually OTM calls. The market prices of these OTM calls clearly reflect greed rather than fear. In this note, we explain why VIX is a volatility index and we also offer several explanations as to why VIX can properly be regarded as a fear gauge.

An overview of this paper is as follows. The next section shows why VIX can properly be regarded as a volatility index. In the following section, we focus on why VIX can also be regarded as a fear gauge. The final section summarizes the paper and offers suggestions for future research.

2. VIX as a volatility index

Market prices of options are widely used to calibrate the stochastic process governing the market price of the underlying asset. When the underlying is a stock index, the stochastic process must respect positivity; a stock index will never hit zero or become negative. One of the easiest ways to respect positivity is to start the index at a positive level and to assume that the worst possible down move in the index over any period is always the same fixed fraction of its beginning of period size. For example, suppose an index starts at 100 and can halve every period, but can never drop by more than half. Then if the worst possible move occurs every period, the index levels are 100, 50, 25, 12.5 etc., so index levels remain positive.

If we couple a constant fractional down move with a constant proportional up move, we obtain a multiplicative binomial process. For example, if the index level can either halve or double each period, then it is following a multiplicative binomial process. For simplicity, consider the further special case of a multiplicative binomial process when the two possible percentage moves differ only in sign. For example, suppose an index is presently at 100 and can only rise or fall by 10% each period. Since the percentage down move is still between 0 and 100%, the index level will always be positive.

For the remainder of this section, we only assume that the S&P500 index level is positive and arbitrage-free. However, we will illustrate all of our results in this context by assuming that the index can only rise or fall by the round figure of 10% each period. Let $S_i > 0$ be the closing level of S&P500 on day $i$. Consider the difference $S_{i+1} - S_i$ between the closing index levels on two adjacent dates. The daily return is defined as the ratio of this difference to the initial index level.

For the remainder of this section, we only assume that the S&P500 index level is positive and arbitrage-free. However, we will illustrate all of our results in this context by assuming that the index can only rise or fall by the round figure of 10% each period. Let $S_i > 0$ be the closing level of S&P500 on day $i$. Consider the difference $S_{i+1} - S_i$ between the closing index levels on two adjacent dates. The daily return is defined as the ratio of this difference to the initial index level. We refer to this random variable as the forward return $f_i \equiv \frac{S_{i+1} - S_i}{S_i}$. In contrast, we define the backward return as the ratio of the difference to the final in-
Index level \( b_i \equiv \frac{S_{i+1} - S_i}{S_{i+1}} \). If time ran backward, then the backward return would become the forward return in the reversed clock. With time running forward, we now show that the backward return is always arithmetically lower\(^1\) than the forward return, i.e. \( b_i \leq f_i \).

If the index rises, then the backward return divides the positive gain by a larger denominator than the forward return, so the fraction is less positive. For example, if the index rises from 100 to 110, then the forward return is \( 10/100 = 0.1 \) or 10%, while the backward return is less positive at \( 10/110 = 0.0909 \) or 9.09%. If the index falls, then the backward return divides the negative gain by a smaller denominator than the forward return, so the fraction is more negative. For example, if the index instead falls from 100 to 90, then the forward return is \(-10/100 = -0.1\), or -10%, while the backward return is more negative at \(-10/90 = -0.1111\) or -11.11%.

The square of the daily return is an estimate of the variance of the daily return. It turns out that the square of the forward return is quite close to the difference between the forward return and the backward return.

Some simple algebra can be used to gain understanding on the magnitude of the approximation error. Since \( S_{i+1} = S_i(1 + f_i) \), the backward return \( b_i \equiv \frac{S_{i+1} - S_i}{S_i} \) can be written in terms of the forward return \( f_i \equiv \frac{S_{i+1} - S_i}{S_{i+1}} \):

\[
b_i = \frac{S_{i+1} - S_i}{S_i(1 + f_i)} = f_i \frac{1}{1 + f_i}.
\]

Now from long division taught in elementary algebra, \( \frac{1}{1 + f} = 1 - f + f^2 - f^3 + \cdots \). As a result:

\[
b_i = f_i - f_i^2 + f_i^3 - f_i^4 + \cdots
\]

It follows that the difference between the forward return and the backward return is just the squared forward return, plus an error whose leading order is a cubed percentage return:

\[
f_i - b_i = f_i^2 + O(f_i^3) \quad \text{(1)}
\]

In continuous time, and for continuous sample path stochastic processes, the cubes and higher powers of returns vanish. As a result, the difference between the forward return and the backward return is the square of the forward return.

The forward return \( f_i \equiv \frac{S_{i+1} - S_i}{S_i} \) can be manufactured by holding \( \frac{1}{S_i} \) units of the index at time \( i \) and borrowing the cost. Assuming no dividends and no interest, the realized gain on this position at time \( i + 1 \) is \( \frac{1}{S_i}(S_{i+1} - S_i) = f_i \), the forward return. How can we manufacture the backward return? The backward return cannot be manufactured by holding \( \frac{1}{S_{i+1}} \) units of the index at time \( i \) because the required holding in the index is anticipating. Fortunately, the gains from a static position in an option can be interpreted as equivalent to the gains from a particular anticipating dynamic trading strategy in its underlying index, even when the volatility process is unknown. For example, consider the gain in intrinsic value \( |S_{i+1} - S_i| \) that arises from being long one ATM straddle at time \( i \). Regardless of the \( S \) dynamics, this gain can always be represented as \( N^a(S_{i+1} - S_i) \), where the number of shares held from time \( i \) to time \( i + 1 \) is \( N_i = 1(S_{i+1} \geq S_i) - 1(S_{i+1} < S_i) \). In words, the strategy is long one share if \( S \) rises and short one share if \( S \) falls. The superscript \( a \) on \( N_i \) indicates that this trading strategy is anticipating since the holdings at period \( i \) clearly depend on \( S_{i+1} \). This example illustrates that a static position in an option is equivalent to a dynamic trading strategy in its underlying which in general is anticipating.

To understand the option position needed to access the backward return, we first recall from calculus that any differentiable function of \( S \) can be treated as the area under the plot of its derivative against \( S \). Let \( A(S) \) be a given differentiable function of \( S \) and consider the plot of \( A'(S) \) against \( S \). As \( S \) moves, \( A \) moves. Let \( \Delta S_i = S_{i+1} - S_i \) be the change in \( S \) and let \( \Delta A_i = A_{i+1} - A_i \) be the corresponding change in \( A \). Using the trapezoidal rule:

\[
\Delta A_i = \left[ \frac{1}{2} A'(S_i) + \frac{1}{2} A'(S_{i+1}) \right] \Delta S_i + e, \quad \text{(2)}
\]

where the error term is:

\[
e = -\frac{(\Delta S_i)^3}{12} A''(m_1), \quad \text{(3)}
\]

\(^1\)As a mnemonic, notice that the letter \( b \) comes before the letter \( f \) in the alphabet.
with $m_t$ between $S_t$ and $S_{t+1}$. In words, (2) says that the change in the area under a function such as $A'(S)$ is approximated by the product of the average height of the function and its base. From (3), the leading term in the error is the cube of the index change.

To illustrate, suppose $A(S) = \ln S$. Then since $A'(S) = \frac{1}{S}$:

$$\Delta(\ln S)_t = \left[ \frac{1}{2} \frac{1}{S_t} + \frac{1}{2} \frac{1}{S_{t+1}} \right] \Delta S_t + O(f_i^3). \quad (4)$$

The LHS is just the log price relative $\Delta(\ln S)_t = \ln(\frac{S_{t+1}}{S_t}) \equiv c_t$, which is the continuously compounded rate of return. Hence, (4) shows that the continuously compounded rate of return is approximated by the blended return that arises by averaging the forward return with the backward return:

$$c_t = \frac{1}{2} f_t + \frac{1}{2} b_t + O(f_i^3). \quad (5)$$

Since $b_t \leq f_t$, we obviously have $b_t \leq c_t \leq f_t$ when the $O(f_i^3)$ term can be ignored. In words, the continuously compounded return $c_t$ is always between the backward return $b_t$ and the forward return $f_t$. From (5), the leading term in the error in treating $c_t$ as a simple average of $b_t$ and $f_t$ is the cube of the return, $f_i^3$, which is typically quite small.

To see how small this approximation error is in practice, let’s examine the error when the daily percentage moves are $\pm 10\%$, which corresponds to a huge annualized volatility of $\sqrt{252} \times 0.1 = 158\%$. If the index rises from 100 to 110, then the continuously compounded non-annualized return is $\ln(1.1) = 0.0953$ or $9.53\%$. The simple average of the forward return of $10\%$ and the backward return of $9.09\%$ is $9.545\%$, which is quite close. If the index instead falls from 100 to 90, then the continuously compounded return is $\ln(0.9) = -0.1053$, or $10.53\%$. The simple average of the forward return of $-10\%$ and the backward return of $-11.11\%$ is $-10.55\%$, which is also quite close.

Now consider the difference between the forward return and the continuously compounded return. Since the continuously compounded return is quite close to the blended return, this difference is quite close to half the difference between the forward return and the backward return which is half the squared return:

$$f_t - c_t = f_t - \left[ \frac{1}{2} f_t + \frac{1}{2} b_t \right] + O(f_i^3)$$

$$= \frac{1}{2} (f_t - b_t) + O(f_i^3)$$

$$= \frac{1}{2} f_t^2 + O(f_i^3). \quad (6)$$

from (1). It follows that if we double the difference between the forward return and the continuously compounded return, we approximate the squared return:

$$2(f_t - c_t) = f_t^2 + O(f_i^3). \quad (7)$$

This observation lies at the heart of the following VIX construction.

Summing over 30 days implies:

$$\sum_{i=0}^{29} 2(f_t - c_t) = \sum_{i=0}^{29} f_t^2 + \sum_{i=0}^{29} O(f_i^3). \quad (8)$$

The LHS contains the continuously compounded return $c_i = \ln S_{i+1} - \ln S_i$. The summing causes telescoping so that:

$$-2 \ln S_{30} + 2 \ln S_0 + \sum_{i=0}^{29} 2 f_t$$

$$= \sum_{i=0}^{29} f_t^2 + \sum_{i=0}^{29} O(f_i^3). \quad (9)$$

Solving for the sum of squared returns:

$$\sum_{i=0}^{29} f_t^2 \approx -2 \ln(S_{30}/S_0)$$

$$+ \sum_{i=0}^{29} \frac{2}{S_i} (S_{i+1} - S_i), \quad (10)$$

since $f_t = \frac{S_{i+1} - S_i}{S_i}$. The first term on the RHS is path-independent, so can be spanned by a static position in 30 day OTM options. Carr and Madan[2] show:

$$-2 \ln(S_{30}/S_0)$$

$$= -\frac{2}{S_0} (S_{30} - S_0) + \int_0^{S_0} \frac{2}{K^2} (K - S_{30})^+ dK$$

---

2 As a mnemonic, notice that the letter $c$ is between the letter $b$ and the letter $f$ in the alphabet.
Substituting (11) in (10) and multiplying by \( \frac{252}{30} \) implies that an estimate of the annualized daily variance rate is:

\[
\frac{252}{30} \sum_{i=0}^{29} f_i^2 \\
\approx \int_{S_0}^{S_0} \frac{252}{15K^2} (K - S_{30})^+ dK \\
+ \int_{S_0}^{\infty} \frac{252}{15K^2} (S_{30} - K)^+ dK \\
+ \sum_{i=0}^{29} \frac{252}{15} \frac{1}{S_i} \left[ \frac{1}{S_0} (S_{i+1} - S_i) \right].
\] (12)

In words, an estimate of the annualized daily variance is approximated by the P&L arising from combining a static position in OTM puts and calls with daily dynamic trading in the index.

Under zero interest rates and dividends, the last term on the RHS of (12) can be synthesized at zero cost by holding \( \frac{252}{30} \left( \frac{1}{S_i} - \frac{1}{S_0} \right) \) units of the index and borrowing the cost. In contrast, the first two terms on the RHS of (12) can be synthesized at zero cost by holding \( 252 \) units of the index and borrowing

\[
\int_{S_0}^{S_0} \frac{252}{15K^2} (K - S_{30})^+ dK \\
+ \int_{S_0}^{\infty} \frac{252}{15K^2} (S_{30} - K)^+ dK
\]

This is very close to the CBOE formula for VIX²:

\[
\text{VIX}^2 = \frac{2}{T} \sum_{i} \frac{\Delta K_i}{K_i^2} e^{RT} Q(K_i) - \frac{1}{T} \left[ \frac{F}{K_0} - 1 \right]^2,
\]

where \( K_0 \) is the first strike below the index level and \( K_i \) is the strike price of the \( i \)th out-of-the-money option; a call if \( K_i > K_0 \); and a put if \( K_i < K_0 \); both put and call if \( K_i = K_0 \) and \( Q(K_i) \) is the midpoint of the bid-ask spread for each option with strike \( K_i \).

It differs from VIX² for several reasons:

1. We have assumed that \( R = 0 \).
2. VIX² uses options of two maturities and then linearly interpolates to achieve a 30 day expiration. As long as the convexity of the term structure of variance rates is small, the impact of this difference is small.
3. VIX² separates puts from calls using the initial forward at their common maturity rather than the initial spot. As long as short term interest rates and short term dividends are small, the impact of this difference is small.
4. VIX² uses discrete strikes rather than a continuum of strikes. As long as there is a fine grid and a wide range of S&P500 strikes trading, the impact of this difference is small.

This section has shown why VIX can properly be regarded as a volatility index. In the next section, we focus on why VIX can also be regarded as a fear gauge.

3. VIX as a fear gauge

There are a few obvious reasons why VIX can also be regarded as a fear gauge. First, the market as a whole is usually considered to be long the S&P500 and risk-averse. Suppose that the S&P500 moves up and down relative to its expected level by the same percentage amount of this level. When the magnitude of these daily moves increases from e.g. ±1% to ±2%, then VIX obviously rises and the utility of risk-averse market participants falls. Alternatively, if the daily moves stay constant at ±1% but the aversion to these moves increases, then VIX might well rise and again the utility of risk-averse market participants falls. Hence, if fear is interpreted as the lowering of expected utility for a risk-averse agent, then a rise in VIX leads to fear.

Suppose now that all investors are risk-neutral, but we continue to assume that the market as a whole is long the S&P500. There are a couple of other reasons why VIX can still be regarded as a fear gauge. It is well known that VIX²T is well approximated by the cost of creating the path-independent payoff \(-2 \ln(S_T/S_0)\). We consider the delta of this position to be negative, even though the Black Scholes value of this payoff is independent of \( S_0 \). The apparent contradiction is re-
solved by recognizing that the payoff depends on $S_0$. As a result, the definition of delta as a hedge ratio and as a first derivative diverge. We treat delta as fundamentally a hedge ratio. Given this definition, delta can be calculated as a first derivative provided the contribution from the payoff function is ignored. As a result, the delta of the claim paying off $-2 \ln(S_T/S_0)$ in negative. This can be regarded as a rationale for treating VIX as a fear gauge.

There is a third reason why VIX can be regarded as a fear gauge. It is well known that the realized variance of S&P500 has historically been negatively correlated with the level of S&P500. Hence, realizations of S&P500 return variance above what was expected ex ante tend to be accompanied by realizations of S&P500 returns below what was expected ex ante. Finance academics often refer to this phenomenon as a leverage effect. It is also well known that the implied variance of index options are positively correlated with subsequent realized variance.

Consider a classical stochastic volatility model:

$$dS_t = \sqrt{V_t} S_t dW_t, \quad t \geq 0,$$  \hspace{1cm} (13)

where $W$ is a $Q$ standard Brownian motion, $S$ is always positive, and $V$ is an unknown stochastic process. While VIX$^2$ is usually defined via option prices, Carr and Lee [1] prove in this context that VIX$^2$ is approximated by a Gaussian weighted average of appropriately defined implied variance rates:

$$VIX^2 T \approx E_Q^0 \int_0^T V_t \, dt$$

$$= \int_{-\infty}^{\infty} e^{\frac{-z^2}{2}} I^2(z) T \, dI.$$  \hspace{1cm} (14)

In (14), the $I^2(z)$ denotes the Black implied variance rate as a function of $z = \frac{d_S X_T - k}{\ln S_0 X_T}$, where $X_T = \ln(F_T/F_0)$, $K = \ln(K/F_0)$, and $B$ is the forward measure in the Black model. One can think of $z$ as the $d_2$ variable in the standard way of writing the Black-Scholes formula for a call.

The typical graph of $I^2(z)$ is downward sloping and mildly convex in $z \in \mathbb{R}$. A standard heuristic is that the ATM implied variance $I(0)$ captures the expected value of future realized variance, the ATM slope $I'(0)$ captures the covariance of future realized variance with returns, and the ATM curvature $I''(0)$ captures the variance of future realized variance. An increase in $I(0)$, or $|I'(0)|$ or $I''(0)$ tends to increase the integral in (14) and hence VIX$^2$. To the extent that a spike up in VIX$^2$ reflects an increase in the forecasted average level of realized variances, the leverage effect suggests that this VIX increase will be followed by lower returns.

However, suppose that risk-neutral investors knew for certain that realized variance of S&P500 evolves independently of the level of S&P500. Is there still a reason that VIX might be called a fear gauge? The Gaussian in (14) is symmetric in $z$, so it appears that the contribution from OTM put implied variances matches the contribution from OTM call implied variances. However, it needs to be remembered that $z \equiv 0$ corresponds to the strike price $K = S_0 e^{-\sigma^2 T/2}$, not $S_0$. If $I^2(z)$ is flat and the put call separator is the mean $S_0$, rather than the median $S_0 e^{-\sigma^2 T/2}$, then more than half of the mass is coming from the put implied variances.

To see how much more the at and OTM puts contribute than the at and OTM calls, note that so long as (13) holds:

$$VIX^2 T \approx E_Q^0 \int_0^T V_t \, dt$$

$$= \int_0^{S_0} \frac{2}{K^2} P_0(K) \, dK$$

$$+ \int_{S_0}^\infty \frac{2}{K^2} C_0(K) \, dK.$$  \hspace{1cm} (15)

The RHS is the exact initial cost of synthesizing a hypothetical variance swap paying $(\ln S)_t = \int_0^T V_t \, dt$ at its maturity date $T$.

Suppose that the constant volatility Black model is holding so that all of the implied variance rates are flat at $\sigma^2$. The cost of creating the floating leg of the variance swap requires a positive investment in both OTM puts and OTM calls. Is the dollar investment in OTM puts matched by the dollar investment in OTM calls? We now show that more dollars are invested in OTM puts than calls when the Black model is holding. If we interpret the dollars invested in OTM puts as due to fear and further interpret the dollars invested in OTM calls as due to greed, then VIX is more about fear than greed. When the implied volatilities slope down in moneyness as they have since 1987, then even more dollars are spent buying.
OTM puts than OTM calls in synthesizing a variance swap.

Let \( X_T \equiv \ln(S_T/S_0) \) be the non-annualized continuously compounded return over the period \([0, T]\). Recall that the cost of creating the floating leg of a variance swap matches the cost of creating the path-independent payoff \(-2X_T\):

\[
E_{0}^{Q} \int_{0}^{T} V_{t} \, dt = E_{0}^{Q}(-2X_T). \tag{16}
\]

Now the path-independent payoff \(-2X_T\) can be decomposed into the contribution from negative \(X_T\) and the contribution from non-negative \(X_T\):

\[
-2X_T = -2X_T1(X_T < 0) - 2X_T1(X_T \geq 0) = 2(0 - X_T)^+ - 2(X_T - 0)^+. \tag{17}
\]

Substituting (17) in (16) implies that

\[
E_{0}^{Q} \int_{0}^{T} V_{t} \, dt = 2E_{0}^{Q}(0 - X_T)^+ - 2E_{0}^{Q}(X_T - 0)^+. \tag{18}
\]

The results of Carr and Madan imply that:

\[
E_{0}^{Q} 2(0 - X_T)^+ = \frac{2}{S_0}P_0(S_0) + \int_{-\infty}^{S_0} \frac{2}{K^2}P_0(K) \, dK \tag{19}
\]

and:

\[
E_{0}^{Q} -2(X_T - 0)^+ = -\frac{2}{S_0}C_0(S_0) + \int_{S_0}^{\infty} \frac{2}{K^2}C_0(K) \, dK. \tag{20}
\]

Thus, the 2 zero strike puts on \(X_T\) are created from at and OTM puts on \(S_T\). Similarly, the short position in the 2 zero strike calls on \(X_T\) are created from at and OTM calls on \(S_T\).

In the Black Scholes model, it is straightforward to value the 2 zero strike puts on \(X_T\):

\[
2E_{0}^{Q}[0 - X_T]^+ = 2\sigma \sqrt{T}N' \frac{\sigma \sqrt{T}}{2} + \sigma^2 TN - \frac{\sigma \sqrt{T}}{2}. \tag{21}
\]

It is also straightforward to value the short position in the 2 zero strike calls on \(X_T\):

\[
E_{0}^{Q} -2(X_T - 0)^+ = -2\sigma \sqrt{T}N' \frac{\sigma \sqrt{T}}{2} + \sigma^2 TN - \frac{\sigma \sqrt{T}}{2}. \tag{22}
\]

Each position is the sum of two terms. The first term on the RHS of (21) is positive, while the first term on the RHS of (22) is negative and has the same absolute value. Clearly, the first term on the RHS of (21) is larger than its counterpart in (22). The last term on the RHS of (21) is larger than the last term on the RHS of (22), since \(N(\cdot)\) is increasing. As a result, the long position in the two zero strike puts is clearly more valuable than the short position in the two zero strike calls. The sum of the two values is \(\sigma^2 T\), since \(N(\sigma \sqrt{T}) + N(-\sigma \sqrt{T}) = 1\).

4. Summary and future research

VIX is both a volatility index and a fear gauge. It is well known that it arises as a volatility index because \(\text{VIX}^2\) is essentially the cost of replicating the floating leg of a variance swap. In this paper, we argued that the reason that the path-independent payoff \(2(e^{X_T} - 1 - X_T)\) has the same value as the path-dependent payoff \((X_T)^\dagger = \int_{0}^{T} (dX_t)^2\) is that increments of \(2(e^{X_t} - 1 - X_t)\) capture the difference between the forward and backward returns. When the underlying price process has positive continuous sample paths in continuous time, the difference between the forward and backward returns is just the variance rate of \(X\).

We also gave four reasons why VIX can properly be regarded as a fear gauge. First, assuming that risk-averse investors are long the S&P500, increases in expected variance or risk aversion raise VIX and lower expected utility. Second, we argued that \(\text{VIX}^2 T\) has the same value as the path-independent claim paying \(\ln(S_T/S_0)\) at \(T\) and that the latter claim has negative delta. Third, the leverage effect implies that abnormal
mally high VIX levels tend to be accompanied by ab-
normally low S&P500 levels. Fourth, when the Black
Scholes model is holding, the VIX construction has
more dollars invested in at and OTM puts than in at and
OTM calls. If the at and OTM puts reflect fear while
the at and OTM calls reflect greed, then VIX is more
about fear than greed. When the negative skew is taken
into account, this fear to greed ratio increases. We con-
clude that there are valid reasons for regarding VIX as
both a volatility index and a fear gauge.

Future research should focus on weaker or alterna-
tive sufficient conditions on the risk-neutral index dy-
namics which lead to greater aggregate investment in
at and OTM puts than calls.

Acknowledgement

I am grateful to Maggie Copeland and Charles
Tapiero for comments. They are not responsible for any
errors.

References

financial.050808.114304.
[2] P. Carr and D. Madan, Optimal positioning in derivative se-
curities, Quantitative Finance 1 (2001), 19–37. doi:10.1080/
713665549.