Statistics of VIX Futures and Applications to Trading Volatility Exchange-Traded Products

M. Avellaneda* and A. Papanicolaou†

February 24, 2018

Abstract

We study the dynamics of VIX futures and ETNs/ETFs. We find that contrary to classical commodities, VIX and VIX futures exhibit large volatility and skewness, consistent with the absence of cash-and-carry arbitrage. The constant-maturity futures (CMF) term-structure can be modeled as a stationary stochastic process in which the most likely state is contango with $VIX \approx 12\%$ and a long-term futures price $V^{\infty} \approx 20\%$. We analyze the behavior of ETFs and ETNs based on constant-maturity rolling futures strategies, such as VXX, XIV and VXZ, assuming stationarity and through a multi-factor model calibrated to historical data. We find that buy-and-hold strategies consisting of shorting ETNs that roll long futures, or buying ETNs that roll short futures, will produce theoretically-sure profits if it is assumed that CMFs are stationary and ergodic (see Proposition 3.1). To quantify further, we estimate a 2-factor lognormal model with mean-reverting factors to VIX and CMF historical data from 2011 to 2016. The results confirm the profitability of buy-and-hold strategies, but also indicate that the latter have modest Sharpe ratios, of the order of $SR = 0.5$ or less, and high variability over 1-year horizon simulations. This is due to the surges in VIX and CMF backwardations which are experienced sporadically, but also inevitably, in the volatility futures market.

*Courant Institute of Mathematical Sciences, New York University, 251 Mercer Street, New York, N.Y. 10012-1185 *avellaneda@courant.nyu.edu*

†Department of Finance and Risk Engineering, NYU Tandon School of Engineering, 6 MetroTech Center, Brooklyn, NY 11201 *ap1345@nyu.edu*

1 The Volatility Index: stylized facts

Volatility futures and exchange-traded products (ETPs) are widely traded instruments in today's markets. VIX-related exchange-traded funds (ETFs) and exchange-traded notes (ETNs), launched in early 2009, have experienced exponential growth in trading volumes. There appears to be a strong appetite by institutional investors for volatility products, for the purpose of managing tail-risk and hedging, as well as trading S&P500 option implied volatility as a stand-alone asset class.[1]

The Chicago Board Options Exchange Volatility index, VIX, (see [CBOE (2014)]) is equal, at any given time, to the square-root of the value of a basket of short-term S&P500 options, in which the portfolio weights depend on the moneyness and time-to-maturity of each constituent option. The portfolio's value corresponds to the fixed leg in a 30-day variance swap on the S&P 500 Index (for details, see e.g., [Whaley (1993)], [CBOE (2014)]). In 2006, the CBOE launched VIX futures contracts, which settle on the spot VIX and provide a means of trading volatility as a commodity.

There is no well-accepted way to "carry" or to "store" VIX for delivery against a futures contract, as one may attempt to do with index or commodity futures. The main reason for this is that the weights of the VIX basket are not known before the settlement date. Therefore, "carrying" VIX would imply dynamic trading ("rolling") at-the-money-forward and near-the-money, constant maturity, 30-day options. The cost of such strategies cannot be calculated accurately in advance.[2] In the event of a sudden market shock, or the perception of an impending shock, the demand for S&P 500 puts, is met by option sellers, and hence will raise option-implied volatilities. The lack of "storage" also implies that, once the equity market's demand for protection subsides, implied volatilities will drop likewise. The conclusion is that VIX should be much more volatile than commodities such as crude oil or equity indices.[3]

The top left plot of Figure 1 displays the VIX time series from 2004 through 2016. Noticeable

[1] A similar exchange-traded volatility market exists in Europe around the $VSTOXX$ index, which tracks the short-term implied volatility of STOXX50, in Deutsche-Borse-XETRA.

[2] An additional complication for replicating VIX with options stems from the fact that VIX is defined as the *square-root* of the value of the basket, but we shall not discuss this further.

[3] Certain commodities, such as electricity or natural gas, are technically difficult to distribute and store. Probably for these reasons they give rise to VIX-like price spikes. On the other hand, crude oil, for example, benefits from ample storage and transportation infrastructures, so that its price volatility is primarily linked to major changes in production/delivery.

spikes in the time series are seen in October 2008 (Lehman Brothers collapse), June 2010 (Greek debt crisis), August 2011 (U.S downgrade by Standard and Poor's), and August 2015 (Reminbi devaluation). We find, empirically, that the time-series of VIX behaves like a stationary process. This statement can be justified statistically, to some extent, using VIX end-of-day historical values for the index. If we use the period from January 1, 1990 to May 31, 2017 (6906 daily observations), Matlab's version of the Augmented Dickey-Fuller test (*adftest*) rejects the presence of a unit root, returning a DFstat=-3.0357 with critical value CV= -1.9416 and pvalue=0.0031. We point out, however, that *adftest* does not reject unit-roots over other time windows, particularly if the events of 2008 are removed from the data. We also studied the stationarity of VIX by a different method, which does not involve the parametric assumptions made in Dickey-Fuller, by splitting the VIX time series into two samples and performing a 2-sample Kolmogorov-Smirnoff (KS) test on the daily differences in VIX. Figure 1 shows the results of the KS tests in support of stationarity, the histogram of VIX, and a power-law fit to the right-tail distribution. The empirical mean is 19.1, the mode is 11.98, and the right-tail is fitted to a Pareto tail with exponent $\alpha = -3.4$. This shows that, as a random variable, VIX is definitely skewed to the right with a heavy tail. On the other hand, VIX has never closed below 9 as of this writing. Once again, we find that we cannot reject a unit root if the 2008 crisis is included in the portion of data before the split. Thus, statistically speaking, we may reject a unit root in VIX if we include the 2008 period in the sample, but we also find that tests conducted over smaller time windows may not be conclusive.

Taking into account the well-known theory of volatility risk-premium (see [Carr & Wu (2009)]) we expect that typically long-dated futures trade higher than the long-term statistical average of VIX. On the other hand, the *most likely level* of the VIX empirical distribution is significantly lower than the average. This is consistent with the fact the futures term-structure is typically upward-sloping, or in contango. The presence of heavy tails for the VIX distribution also indicates that, sporadically, spot VIX will be higher than its long-term average, so there are periods of backwardation of the futures curve, either partly (only for front-month contracts) or fully across the entire term-structure. These periods of backwardation can be seen as VIX surges rapidly, driven by exogenous news and the increase in short-term option premia.

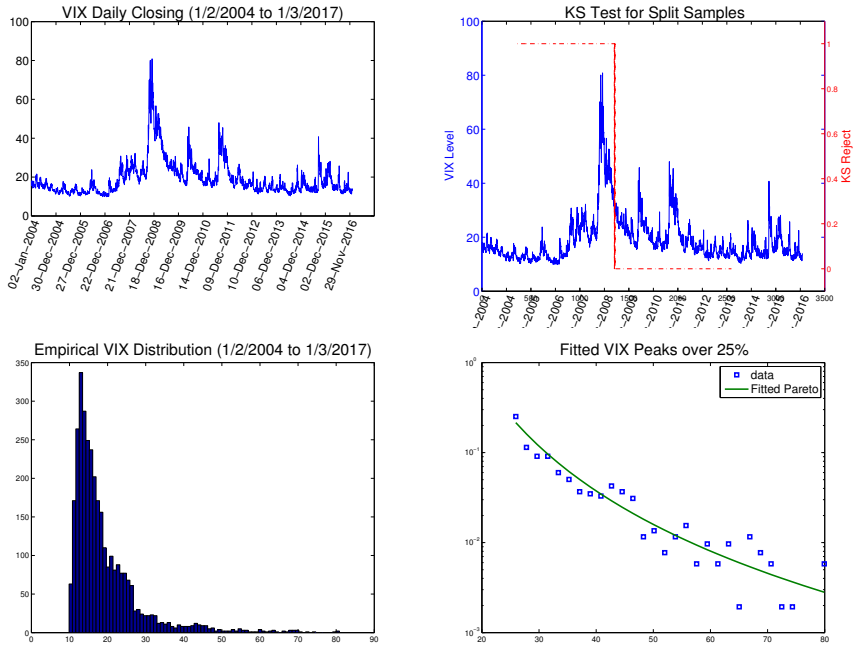

Figure 1: **Top Left:** The time series of VIX, from 2004 through 2016. **Top Right:** Two-sample Kolmogorov-Smirnoff (KS) test for the daily changes of VIX, with different splitting dates in the data. If the 2008 crisis is included in the left data sample, then KS does not reject. **Bottom Left:** Empirical distribution from VIX time series. Mean=19.1 and mode=11.98. **Bottom Right:** Fit of a Pareto density $p(v) = c(v - v_0)^a$ to the empirical distribution of VIX values over 25 (vertical axis in log scale). Estimated parameters are $v_0 = 5$, $a = -3.4$, and $c = 8.8132$.

2 VIX Futures

Currently, VIX futures have 9 open monthly expirations as well as six additional contracts settling weekly in the front of the curve.[4] VIX futures provide a straightforward approach for market participants to position in 30-day forward near-the-money volatility corresponding to each of the settlement dates.[5]

To capture the evolution of the VIX futures term-structure, it is convenient to work with

[4] This article will concern the monthly contracts, where the bulk of the liquidity is concentrated.

[5] VIX futures settle in the morning of the Wednesday of the week of the third Friday of the corresponding settlement month.

constant-maturity futures (CMFs).[6] Accordingly, we define V_t^τ as the price of a theoretical futures contract with maturity (or tenor) τ observed at date t. In practice, V_t^τ is defined as a linear interpolation between the two successive monthly contracts which expire immediately before and immediately after τ years, i.e. such that their maturities "bracket" τ. If τ is shorter than the maturity of the first monthly futures, V^τ is defined as the linear interpolation between spot VIX and the price of the first contract. Denoting by F_1 and F_2 the prices of the monthly contracts (or VIX and the first contract) which have maturities τ_1 and τ_2, $\tau_1 \leq \tau < \tau_2$, we have

$$V^\tau = \frac{\tau_2 - \tau}{\tau_2 - \tau_1} F_1 + \frac{\tau - \tau_1}{\tau_2 - \tau_1} F_2 \ . \tag{1}$$

If $F_1 = VIX$ then $\tau_1 = 0$.[7] With this approach, we can build the constant-maturity futures curve from the available data on each trading date and perform statistical analysis on the daily observations of the curves.

Our main modeling assumption is that the time-series of CMFs is stationary, i.e., for each τ the time series V_t^τ for $0 < t < \infty$ can be modeled as stationary process.

2.1 Principal Component Analysis

We considered the time series of daily settlement prices from February 8, 2011 to December 15, 2016 for the 9 open monthly VIX futures contracts, which have ticker symbols VX1, VX2, VX3,...., VX9, as well as the (spot) VIX end-of-day value corresponding to each date. We discard the 9th monthly contract, after we found that VX9 is not liquid and that the settlement prices tend to be noisy and often missing. This leaves us with data on VIX and the first 8 monthly contracts. From this data, we constructed 8 CMFs using the linear weights as in (1), corresponding to the maturities $\tau_j = j\frac{30}{365}$ with $j = 0, 1, 2 \ldots, 7$, spanning from VIX ($\tau = 0$) to 8 months. There are $N = 1,499$ observed dates.

[6] This is entirely similar to the concept of Constant Maturity Treasuries [H.15], see also [Alexander & Korovalis(2013)] who first studied it for VIX futures.

[7] Clearly, linear interpolation is a possible choice but by no means the only one. Linear interpolation shall be replaced by a smooth function of τ later on, when we introduce a parametric model for the term-structure, but there is little difference in practice, except for the somewhat annoying fact that the derivative with respect to τ has jumps in the linear interpolation model.

We considered the $N \times 8$ data matrix

$$\mathbf{V}_{ij} = ln(V_{t_i}^{\tau_j}) - \overline{ln(V^{\tau_j})} \ , \ i = 1, \dots N, \quad j = 0, \dots 7,$$

where

$$\overline{ln(V^{\tau_j})} = \frac{1}{N} \sum_{i=1}^{N} ln(V_{t_i}^{\tau_j}) \ .$$

We performed a singular-value decomposition of the data matrix. The singular value decomposition (SVD) decomposes the centered matrix,

$$\mathbf{V} = US\psi' \ ,$$

where U is an $N \times 8$ matrix with orthonormal columns, S is an 8×8 diagonal matrix containing the singular values, and ψ is an 8×8 orthonormal matrix whose columns are the principal components used to form any given futures curve. In other words, we have

$$ln(V_{t_i}^{\tau_j}) = \overline{ln(V^{\tau_j})} + \sum_{\ell=1}^{8} a_{i\ell}\psi_{j\ell} \ ,$$

where the coefficient matrix is $a = US$. Figure 2 shows the results of the PCA.

We note that this analysis differs from [Alexander & Korovalis(2013)], where the authors used PCA to analyze 1-day log-returns; see also [Litterman & Scheinkman (1991)], [Laloux *et al.* (2000)], and [Dobi (2014)]. In contrast, we consider here the fluctuations of the VIX futures curve around its *long-term average*. This is consistent with the mean-reversion/stationarity assumption, and the *ansatz* for the log CMFs described above. In other words, this approach to PCA is intended to determine various curves which are needed to "reconstruct" the CMF curve at any particular date, viewed as a superposition of shapes. We call this the *Equilibrium PCA* to distinguish it from the analysis of [Alexander & Korovalis(2013)]. Equilibrium PCA provides an identikit of the VX curve, in the spirit of image recognition: each component is seen as providing a "shape" or feature which is superimposed to the average, with a corresponding loading coefficient. Each curve in the data

would correspond to a set of loadings applied to the different principal components.

We found that the first principal component accounts for 72% of the term-structure shape, adding the second component increases it to 90%, adding the third increases it to 96%, and fourth increases it to 97%. Figures 3 and 4 show poor quality for selected dates of fit when having only one component and a significant improvement from inclusion of a second. Hence, it is reasonable to consider a parsimonious model with $d = 2$:

$$ln(V_{t_i}^{\tau_j}) = \overline{ln(V^{\tau_j})} + \sum_{\ell=1}^{2} a_{i\ell}\psi_{j\ell} + \varepsilon_i \ ,$$

where ε_i is a noise term.

Using the Equilibrium PCA we determine the *most likely CMF curve given the data*. To find it, we consider the time series of the loadings on the different shape factors $\{a_{i\ell},\ i = 1, 2, ..., N\}$ for each ℓ and calculate the mode, or most frequently observed value. Assuming independence (or weak dependence) between the factors, we postulate that the most likely CMF curve is

$$mode\Big(ln(V_{t_i})\Big) \approx \overline{ln(V)} + mode\Big(a_{i1}\Big)\psi_1 + mode\Big(a_{i2}\Big)\psi_2 \ .$$

The top plots in Figure 5 show the histograms of a_{i1} and a_{i2}, and the bottom plot shows the most-likely curve, which is clearly in contango.

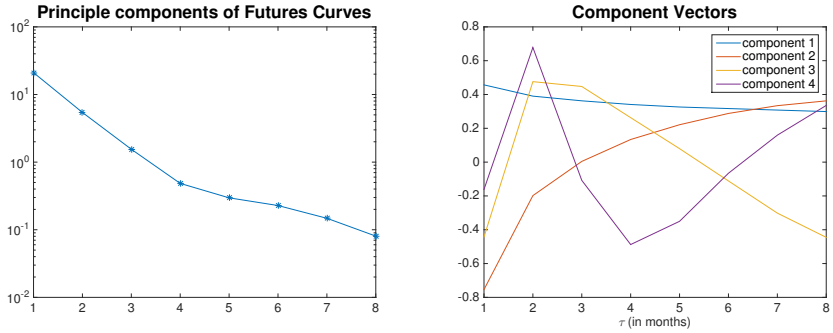

Figure 2: **Left:** For the CMFs, the singular values from the SVD's diagonal matrix S shown on log scale. The first singular component alone accounts for 72% of variation, adding the second component increases it to 90%, adding the third increases it to 96%, and fourth increases it to 97%. **Right:** The first four singular vectors; the "zero-th"-order component (not shown) is just the overall average level $\overline{ln(V^\tau)} = \frac{1}{N}\sum_i ln(V_{t_i}^\tau)$.

3 Rolling Futures and Exchange-Traded Notes

In a typical rolling futures strategy, positions are transferred from one expiration to another as contracts expire. Perhaps the simplest example is investing in a long position in the first contract, reinvesting all profit/losses in futures and, on the expiration date of the contract, close the long position and simultaneously open a long position for the same dollar amount in the second contract.[8]

3.1 Rolling the front-month contract

We denote the futures contract maturing at date T_1 by F_{t,T_1}, where t is the current date. We consider the situation of a fund that buys a number of contracts with notional value equal to the equity in the fund, rebalances daily investing the profit/loss in futures, and rolls the entire position into the next contract at the expiration date of the current contract.[9] We assume that the cash in the fund is remunerated at rate r and neglect futures margin costs and other expenses. The total

[8] VIX futures settle in the morning of the contract's expiration date. If we use daily close-to-close data it is somewhat simpler, to roll the position at the close of the day *prior* to the futures' expiration date. One could also let the first contract expire and open simultaneously a position in the second contract. However, the latter strategy would require intraday data for backtesting purposes. We also reinvest profits, which is consistent with the way ETN indices are constructed.

[9] Rebalancing is not necessary, but it is consistent with how an ETP operates: the notional amount invested in futures is equal to the total equity of the fund at any given time.

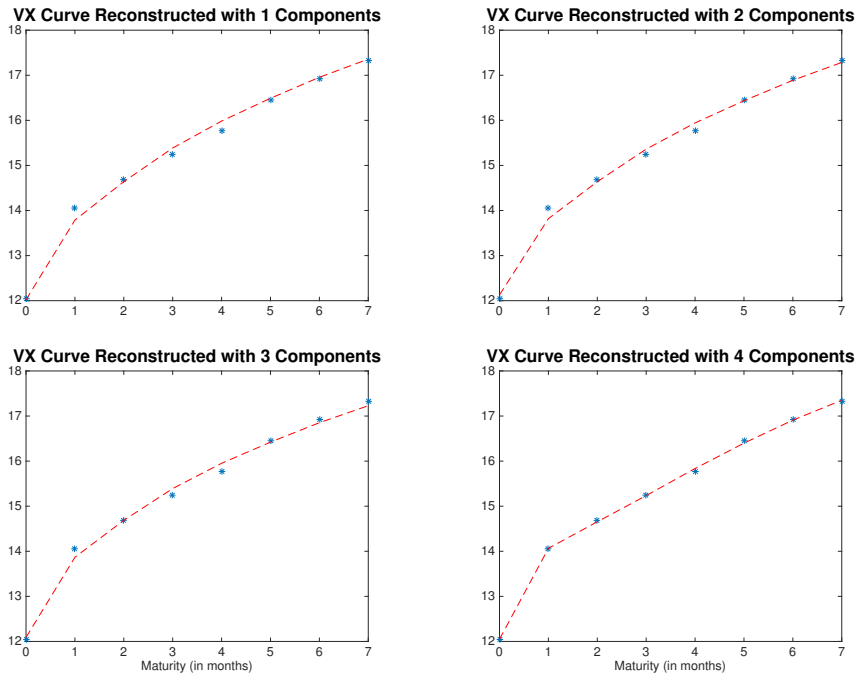

Figure 3: PCA reconstruction of VX futures curve for April 10th 2014. PCA calculated from daily VX curves from February 8th 2011 to December 15th 2016.

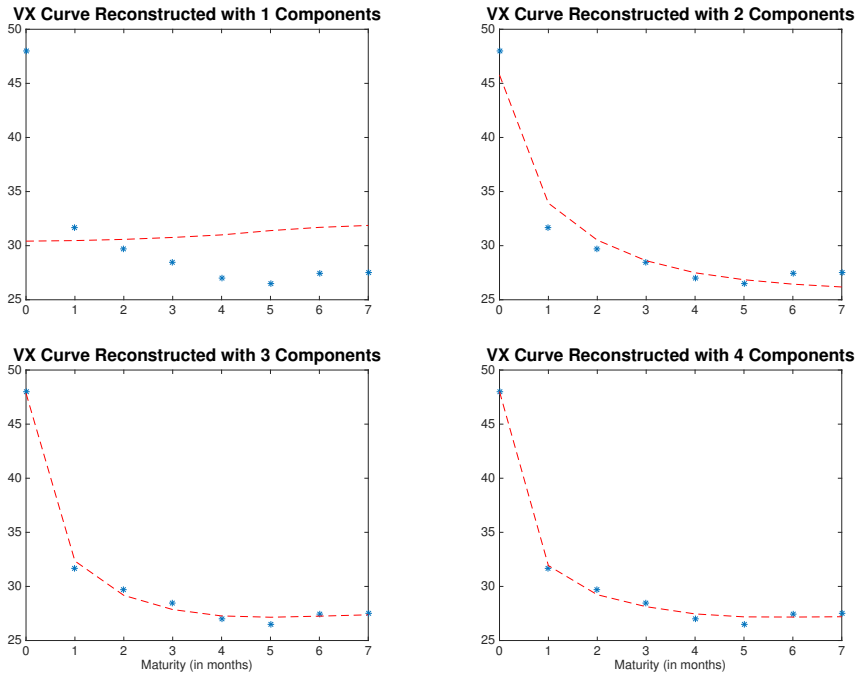

Figure 4: PCA reconstruction of VX futures curve for August 8th 2011. These curves are not those of the VIX's most-likely state, but rather a tumultuous period around the first-ever downgrade of U.S. government debt.

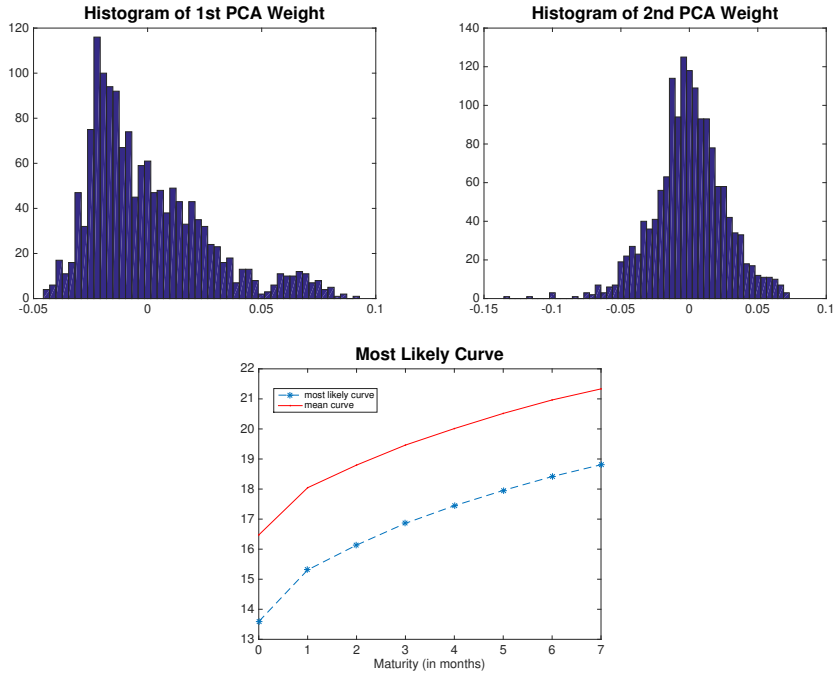

Figure 5: **Top Plots:** The histogram of weights for the 1st and 2nd principal components, a_{i1} and a_{i2} respectively. **Bottom Plot:** Recall mean VIX higher than mode VIX. Here mean curve is higher than mode curve, $mode\Big(ln(V_{t_i})\Big) \approx \overline{ln(V)} + mode\Big(a_{i1}\Big)\psi_1 + mode\Big(a_{i2}\Big)\psi_2$.

value of the fund, I_t, satisfies

$$\frac{dI_t}{I_t} = \begin{cases} \frac{dF_{t,T_1}}{F_{t,T_1}} + rdt & \text{if } t < T_2 - dt \\ \frac{dF_{t,T_2}}{F_{t,T_2}} + rdt & \text{if } t \geq T_2 - dt \ . \end{cases}$$

Here, T_1 is the expiration date of the front contract, T_2 is the expiration date of the second contract, r is the interest rate, and $dX_t = X_{t+dt} - X_{dt}$ is the one-day change from date t to $t + dt$. We shall refer to I_t as the *total return index* or TRI of the strategy.

It is convenient to define the function

$$\phi(t) = \begin{cases} F_{t,T_1} & \text{if } t < T_2 - dt \\ F_{t,T_2} & \text{if } t = T_2 - dt \ , \end{cases} \tag{2}$$

which represents the settlement price of the front-month contract at time t. Writing

$$\phi(t) = b(t)F_{t,T_1} + (1 - b(t))F_{t,T_2} \ , \tag{3}$$

where $b(t) = 1$ if $t < T_2$ and $b(t) = 0$ if $t = T_2$, we have

$$d\phi(t) = b(t)dF_{t,_1} + (1 - b(t))dF_{t,T_2} + (F_{t,T_2} - F_{t,T_1})\delta(t - T_2)dt \ ,$$

where $\delta(\cdot)$ is the Dirac delta. Therefore, from equation (2) we have

$$\frac{d\phi(t)}{\phi(t)} = \frac{dI_t}{I_t} - rdt + \frac{(F_{t,T_2} - F_{t,T_1})}{\phi(t)}\delta(t - T_2)dt \ ,$$

or

$$\frac{dI_t}{I_t} = \frac{d\phi(t)}{\phi(t)} - \frac{(F_{t,T_2} - F_{t,T_1})}{\phi(t)}\delta(t - T_2)dt + rdt \ .$$

Suppose that the front-month contract price follows a stationary distribution. Then, heuristically,

if the curve is in contango, rolling a long position will lose money on average and, conversely it will earn money when the curve is in backwardation.[10] Similar considerations apply to other rolling strategies. For instance, one could envision a strategy which rolls the futures in the middle of the month, i.e. when $\theta + t = \frac{T_1+T_2}{2}$.[11]

3.2 Continuous (daily) rolling

Consider a yet-to-be-determined, decreasing function $b(t)$ such that $b(T_1+) = 1$ and $b(T_2) = 0$, where $b(t)$ represents the fraction of the total number of contracts invested in the front contract, and $1-b(t)$ is the fraction invested in the second-month contract. The (contract-weighted) average maturity is therefore

$$\theta(t) = b(t)(T_1 - t) + (1 - b(t))(T_2 - t) \ .$$

If we set $\theta(t) = \theta =$ constant, we find that $b(t) = \frac{T_2-t-\theta}{T_2-T_1}$ and $\dot{b}(t) = -\frac{1}{T_2-T_1}$. According to the definition of CMF futures given above, we have

$$\phi(t) = b(t)F_{t,T_1} + (1 - b(t))F_{t,T_2} = V_t^\theta \ ,$$

whence

$$dV_t^\theta = b(t)\ dF_{t,T_1} + (1 - b(t))\ F_{t,T_2} + (F_{t,T_2} - F_{t,T_1})\frac{dt}{T_2 - T_1} \ .$$

In particular, setting

$$w_1(t) = \frac{b(t)F_{t,T_1}}{V_t^\theta} \quad , \quad w_2(t) = \frac{(1 - b(t))F_{t,T_2}}{V_t^\theta} \ ,$$

[10] This is clear if the processes are discrete or have finite variation. The assertion is also correct in the case of continuous-time diffusions, as we show hereafter

[11] Mid-month rolling is essentially the strategy used by the US Oil ETN, USO.

we obtain

$$\begin{aligned}\frac{dV_t^\theta}{V_t^\theta} &= w_1(t)\frac{dF_{t,T_1}}{F_{t,T_1}} + w_2(t)\frac{dF_{t,T_2}}{F_{t,T_2}} + \frac{F_{t,T_2} - F_{t,T_1}}{V_t^\theta}\frac{dt}{T_2 - T} \\ &= w_1(t)\frac{dF_{t,T_1}}{F_{t,T_1}} + w_2(t)\frac{dF_{t,T_2}}{F_{t,T_2}} + \Big[\frac{\partial ln(V_t^\tau)}{\partial \tau}\Big]_{\tau=\theta} dt \ .\end{aligned}$$

It follows that the TRI of a continuous rolling long futures strategy with average constant maturity θ satisfies

$$\begin{aligned}\frac{dI_t^\theta}{I_t^\theta} &= w_1(t)\frac{dF_{t,T_1}}{F_{t,T_1}} + w_2(t)\frac{dF_{t,T_2}}{F_{t,T_2}} + rdt \\ &= \frac{dV_t^\theta}{V_t^\theta} - \Big[\frac{\partial ln(V_t^\tau)}{\partial \tau}\Big]_{\tau=\theta} dt + rdt \ , \end{aligned} \tag{4}$$

(see Chapter 7 in [Bergomi (2016)]). The daily return of the index is the sum of the daily percent change of the CMF with maturity θ and a "drift" term which depends on the slope of the CMF curve at $\tau = \theta$. This term is negative if the curve is in contango and positive if the curve is in backwardation at $\tau = \theta$. The weights $w_1(t)$ and $w_2(t)$ correspond to the proportions of assets (dollar-weights) invested in each of the contracts in order to maintain a constant maturity θ.[12]

3.3 Exchange-traded funds and notes

Most, but not all, VIX ETFs and ETNs correspond to constant-maturity rolling strategies.[13] For $\theta = 30$ days, the above example corresponds to the investable index tracked by VXX, the most liquid VIX ETN. Inverse VIX ETNs and ETFs, such as XIV and SVXY correspond to rolling short

[12] Because contract prices are typically different, the weights depend on the futures prices as well as on the times-to-maturity of each contract.

[13] For information about VIX ETNs, see [Alexander & Korovalis(2013)] and [ETNs (2016)]

positions, *i.e.*, have total-return indices,

$$\begin{aligned}\frac{dJ_t^\theta}{J_t^\theta} &= -w_1(t)\frac{dF_{t,T_1}}{F_{t,T_1}} - w_2(t)\frac{dF_{t,T_2}}{F_{t,T_2}} + rdt \\ &= -\frac{dV_t^\theta}{V_t^\theta} + \Big[\frac{\partial ln(V_t^\tau)}{\partial \tau}\Big]_{\tau=\theta} dt + rdt \ . \end{aligned} \tag{5}$$

In the latter case, the exposure to contango is positive: the strategy value has a positive drift if the CMF curve is in contango. Leveraged ETFs, such as $UVXY$ associated with long or short front month continuous rolling are defined similarly.

Another noteworthy example of a VIX ETN, which is not a simple rolling futures strategy, is the VXZ that is constructed using 4 rolling contracts VX4, VX5, VX6 and VX7. It corresponds to the phi-function

$$\phi_2(t) = \frac{b(t)}{3}F_{t,T_4} + \frac{1}{3}F_{t,T_5} + \frac{1}{3}F_{t,T_6} + \frac{1-b(t)}{3}F_{t,T_7} \ ,$$

where $b(t)$ drops linearly from 1 to zero between T_3 and T_4. The corresponding total-return index is defined by

$$\begin{aligned}\frac{dI^{VXZ}}{I_t^{VXZ}} &= \frac{d\phi_2(t)}{\phi_2(t)} - \frac{1}{3\phi_2(t)}\sum_{i=1}^{3}\Big(\frac{F_{t,T_{4+i}} - F_{t,T_{3+i}}}{T_{4+i} - T_{3+i}}\Big)dt + rdt \\ &= \frac{d\phi_2(t)}{\phi_2(t)} - \frac{1}{3\phi_2(t)}\sum_{i=1}^{3} V_t^{\theta_i}\Big[\frac{\partial ln V_t^\tau}{\partial \tau}\Big]_{\tau=\theta_i} dt + rdt \ , \end{aligned}$$

where $\theta_i = 90 + 30 \times i$ days. Notice that can we can also view $\phi_2(t)$ as the average of *three* CMFs, namely

$$\phi_2(t) = \frac{1}{3}\Big(V_t^{\theta_1} + V_t^{\theta_2} + V_t^{\theta_3}\Big) \ .$$

The average maturity of VXZ is $\frac{\theta_1+\theta_2+\theta_3}{3} = (120 + 150 + 180)/3 = 150$ days, *i.e.* 5 months. Table 1 lists some of the VIX ETNs commonly traded in 2017.

ETN	leverage	average maturity (months)	inception date
VXX	1	1m	Jan 2009
VXZ	1	5m	Jan 2009
XIV	-1	1m	Nov 2010
UVXY	2	1m	Oct 2011
ZIV	-1	5m	Nov 2010
TVIZ	2	5m	Nov 2010

Table 1: Some of the VIX ETNs with high trading volume in 2017.

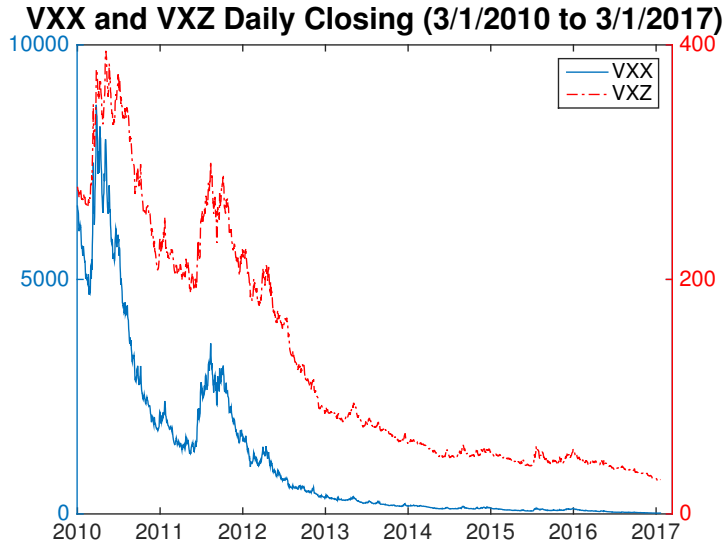

Figure 6: The historical time series of VXX and VXZ prices. The VXX rolls between VX1 and VX2 to maintain a 1-month average maturity, and the VXZ rolls between the VX4, VX5, VX6 and VX7 contracts to maintain an average maturity of 5-months. The VXZ has less volatility than the VXX because it is an average and because it rolls in longer-term contracts; the VXX rolls in 1 and 2-month contracts that are typically more volatile.

3.4 Integrated TRI equations and ETP buy-and-hold strategies

Next, we derive the integrated form of the equations (4) and (5). By Itô's Formula, we have

$$d\ lnV_t^\tau = \frac{dV_t^\tau}{V_t^\tau} - \frac{1}{2}\Big(\frac{dV_t^\tau}{V_t^\tau}\Big)^2, \tag{6}$$

where the last term on the right-hand side corresponds to the quadratic variation process of lnV_t^τ. A similar equation holds for lnI_t^τ. From equation (4) we also deduce that the quadratic variations of both processes are equal:

$$\Big(\frac{dV_t^\tau}{V_t^\tau}\Big)^2 = \Big(\frac{dI_t^\tau}{I_t^\tau}\Big)^2.$$

Thus, equation (4) is equivalent to

$$d\ lnI_t^\tau = d\ lnV_t^\tau - \frac{\partial lnV_t^\tau}{\partial \tau}dt + rdt\ ,$$

or, in integrated form,

$$\frac{I_t^\tau}{I_0^\tau} = e^{rt}\frac{V_t^\tau}{V_0^\tau}\ exp\Big(-\int_0^t \frac{\partial lnV_s^\tau}{\partial \tau}ds\Big)\ . \tag{7}$$

In a similar vein, consider an inverse ETN which tracks a CMF with maturity θ. If we denote this ETN (or its underlying index) by J^θ, we have, from (5),

$$\frac{J_t^\tau}{J_0^\tau} = e^{rt}\frac{V_0^\tau}{V_t^\tau}\ exp\Big(\int_0^t \frac{\partial lnV_s^\tau}{\partial \tau}ds\Big)\ . \tag{8}$$

Equations (7) and (8) display an explicit relation between the total-return indices of ETNs for rolling futures, the constant-maturity futures and the slope of the CMF curve. The derivation of the equations for total-return index of VXZ, as well as other ETNs based on rolling a basket of futures, is left to the reader.

3.5 Long-time behavior of ETNs under stationarity

As an application, consider a trader who sells short an ETN tracking I^τ and invests the proceeds in a money market account paying interest r. The market value of his portfolio at time t in dollars at time zero is ,

$$I_0^\tau - e^{-rt} I_t^\tau = I_0^\tau \left(1 - \frac{V_t^\tau}{V_0^\tau} \, exp\Big(- \int_0^t \frac{\partial ln V_s^\tau}{\partial \tau} ds \Big) \right) .$$

Similarly, a trader who borrows J_0^τ dollars and invests in the inverse ETN with maturity τ will have a present value equal to

$$e^{-rt} \, J_t^\tau - J_0^\tau = J_0^\tau \left(\frac{V_0^\tau}{V_t^\tau} \, exp\Big(\int_0^t \frac{\partial ln V_s^\tau}{\partial \tau} ds \Big) - 1 \right) ,$$

suggesting that the trader will achieve unbounded gains by investing in the inverse VIX ETN in the long run.

If we assume that V_t^τ is ergodic (a mild assumption in this context) then

$$\lim_{t \to \infty} \frac{1}{t} \int_0^t \frac{\partial ln V_s^\tau}{\partial \tau} ds = E\Big[\frac{\partial ln V^\tau}{\partial \tau} \Big] ,$$

where the expectation is taken over the equilibrium measure of the CMF curve process. We conclude that:

Proposition 3.1 *Assume that V_t^τ is a stationary and ergodic process and that $E\Big[\frac{\partial ln V^\tau}{\partial \tau} \Big] > 0$. Then,*

$$\begin{aligned} \lim_{t \to \infty} \Big(I_0^\tau - e^{-rt} I_t^\tau \Big) &= 0 \quad \textit{and} \\ \lim_{t \to \infty} \Big(e^{-rt} J_t^\tau - J_0^\tau \Big) &= \infty \ . \end{aligned} \tag{9}$$

In other words, under our assumptions, shorting the long VIX ETN and going long the short VIX ETN both produce sure profits as $t \to \infty$.

Clearly, these theoretical profits are not free of risk. Equations (7) and (8) show that if V_t^τ is volatile, the random variable $\frac{V_0^\tau}{V_t^\tau}$ can be very small even if t is large. Backwardation associated with high volatility could make $\frac{\partial ln V_s^\tau}{\partial \tau}$ negative as well. Hence the return of the buy-and-hold, long XIV, could be negative over any time interval $[0, t]$.

In order to quantify the profitability, if any, of such strategies in the real world, we resort to a simple parametric model for the evolution of CMFs. This model shall provide further insight on whether of not we can regard CMFs as being a stationary process, and on the performance rolling futures strategies and trading ETPs.

4 A Parametric Model for CMFs

4.1 Lognormal Markov model for VIX

We consider a model in which $ln(VIX_t)$ follows a multi-factor Gaussian diffusion, *i.e.*

$$VIX_t = exp\Big(\sum_{i=1}^{d} X_{it}\Big) \ ,$$

with each factor X_i following a one-dimensional Ornstein-Uhlenbeck (OU) process

$$dX_{it} = \kappa_i \left(\mu_i - X_{it}\right) dt + \sigma_i dW_{it}^P, \ \kappa_i > 0, \quad i = 1, ..., d \ .$$

Here $(W_{1t}^P, ...W_{d\ t}^P)$ is a d-dimensional vector of Brownian motions with correlation matrix ρ_{ij}.

According to this model, VIX is stationary with an equilibrium distribution such that VIX is log-normal, with

$$E[lnVIX] = \sum_{i=1}^{d} \mu_i \ , \qquad var(lnVIX) = \sum_{ij=1}^{d} \frac{\sigma_i \sigma_j \rho_{ij}}{\kappa_i + \kappa_j} \ ,$$

and, accordingly,

$$E[VIX] = exp\Big(\sum_{i=1}^{d} \mu_i \ + \frac{1}{2} \sum_{ij=1}^{d} \frac{\sigma_i \sigma_j \rho_{ij}}{\kappa_i + \kappa_j}\Big) \ .$$

4.2 From VIX to CMFs

To extend this model to CMFs, we assume (see *e.g.* Cox, Ingersoll, Ross [Cox *et al.* (1985)]) that there exists a probability measure Q and a W^P-adapted d-dimensional process λ_t (the market prices of risk or MPRs) such that (i) the processes

$$dW_{it}^Q := dW_{it}^P + \lambda_{it} dt \quad i = 1, ..., d \ ,$$

are standard Brownian motions under Q (with correlation ρ_{ij}), and (ii) every futures price is the expected value under Q of VIX at the corresponding expiration date:

$$V_t^\tau = E_t^Q \Big[VIX_{t+\tau} \Big] = E_t^Q \Big[exp\Big(\sum_{i=1}^{d} X_{i\ t+\tau} \Big) \Big] \ . \tag{10}$$

Following [Cox *et al.* (1985)] and others, we choose the MPRs to be linear functions of the factors,

$$\lambda_{it} = p_i + q_i X_{it} \ ,$$

where p_i and q_i are constants. With this special choice, X_t is distributed as a multi-dimensional OU process under Q, *i.e.,*

$$dX_{it} = \sigma_i dW_t^Q + \overline{\kappa}_i \left(\overline{\mu}_i - X_{it} \right) dt, \quad \text{for} \quad i = 1, ..., d \ ,$$

where

$$\overline{\kappa}_i = \kappa_i + \sigma_i q_i \ ,$$

and

$$\overline{\mu}_i = \frac{\kappa_i \mu_i - \sigma_i p_i}{\kappa_i + \sigma_i q_i} \ ,$$

are new long-term mean and mean-reversion parameters.[14] In particular,

$$X_{it+\tau} = e^{-\overline{\kappa}_i \tau} X_{it} + \left(1 - e^{-\overline{\kappa}_i \tau}\right) \overline{\mu}_i + \sigma_i \int_t^{t+\tau} e^{-\overline{\kappa}_i (t+\tau-s)} dW_s^Q \ .$$

Substituting this expression in equation (10) and computing the conditional expectation, we obtain the CMF curve process under the probability Q, i.e., similar to [Bergomi (2005), Bergromi (2008)] we have,

$$ln\ V_t^\tau = ln\ V^\infty + \sum_i e^{-\overline{\kappa}_i \tau} \left(X_{it} - \overline{\mu}_i\right) - \frac{1}{2} \sum_{ij} \frac{\sigma_i \sigma_j \rho_{ij}}{\overline{\kappa}_i + \overline{\kappa}_j} e^{-(\overline{\kappa}_i + \overline{\kappa}_j)\tau},$$

where $V^\infty = exp\left(\sum_i \overline{\mu}_i + \frac{1}{2} \sum_{ij} \frac{\sigma_i \sigma_j \rho_{ij}}{\overline{\kappa}_i + \overline{\kappa}_j}\right)$ represents the asymptotic value of the CMF curve as $\tau \to \infty$.

A further simplification can be made. Since $X_{it} - \overline{\mu}_i$ has mean zero under Q, we set set $Y_i := X_{it} - \overline{\mu}_i$. Using the relation between the P and Q measures, we derive a stochastic differential equation for Y_i in the P-measure, namely

$$\begin{aligned} dY_i &= dX_i = \sigma_i dW_t^P + \kappa_i(\mu_i - X_i)dt \\ &= \sigma_i dW_t^P + \kappa_i(\mu_i - (Y_i + \overline{\mu}_i)dt \\ &= \sigma_i dW_t^P + \kappa_i(\mu_i^Y - Y_t)dt \ , \end{aligned}$$

with $\mu_i^Y = \mu_i - \overline{\mu}_i$.

The proposed statistical model for the CMF curve dynamics is:

$$V_t^\tau = V^\infty exp\Big(\sum_i e^{-\overline{\kappa}_i \tau} Y_i - \frac{1}{2} \sum_{ij} \frac{\sigma_i \sigma_j \rho_{ij}}{\overline{\kappa}_i + \overline{\kappa}_j} e^{-(\overline{\kappa}_i + \overline{\kappa}_j)\tau}\Big) \ , \tag{11}$$

with

$$dY_i = \sigma_i dW_t^P + \kappa_i(\mu_i^Y - Y_t)dt \ .$$

[14] Notice that there is a one-to-one correspondence between (p_i, q_i) and $(\overline{\mu}_i, \overline{\kappa}_i)$. In other words, the passage from the statistical measure P and the pricing measure Q is tantamount to modifying κ_i and μ_i for each factor.

Remark According to this model, the equilibrium position of the CMF curve is given by

$$V_{eq}^{\tau} = V^{\infty} exp\Big(\sum_i e^{-\overline{\kappa}_i \tau} \mu_i^Y - \frac{1}{2} \sum_{ij} \frac{\sigma_i \sigma_j \rho_{ij}}{\overline{\kappa}_i + \overline{\kappa}_j} e^{-(\overline{\kappa}_i + \overline{\kappa}_j)\tau}\Big) \ .$$

4.3 Profitability of rolling futures strategies

We calculate the theoretical profit/loss of a rolling futures strategy with average maturity τ. From (4), we have

$$\frac{dI^{\tau}}{I^{\tau}} = r dt + \frac{dV^{\tau}}{V^{\tau}} - \frac{\partial ln\ V^{\tau}}{\partial \tau} dt \ .$$

Applying Itô's Formula to (11), we find that

$$\frac{dV^{\tau}}{V^{\tau}} = \sum_i e^{-\overline{\kappa}_i \tau} dY_i + \frac{1}{2} \sum_{ij} \sigma_i \sigma_j \rho_{ij} e^{-(\overline{\kappa}_i + \overline{\kappa}_j)\tau} \ dt \ .$$

Furthermore, differentiation of $ln\ V^{\tau}$ with respect to τ in (11) yields

$$\frac{\partial ln\ V^{\tau}}{\partial \tau} = -\sum_i \overline{\kappa}_i \ e^{-\overline{\kappa}_i \tau} Y_i + \frac{1}{2} \sum_{ij} \sigma_i \sigma_j \rho_{ij} e^{-(\overline{\kappa}_i + \overline{\kappa}_j)\tau} \ .$$

Subtracting one expression from the other and rearranging,

$$\begin{aligned}
\frac{dI^{\tau}}{I^{\tau}} &= r dt + \sum_i e^{-\overline{\kappa}_i \tau} \left(dY_{it} + \overline{\kappa_i}\, Y_{it} dt\right) \\
&= r dt + \sum_i e^{-\overline{\kappa}_i \tau} \left(dY_i + \kappa_i\, Y_{it} dt + (\overline{\kappa_i} - \kappa_i) Y_{it} dt\right) \\
&= r dt + \sum_i e^{-\overline{\kappa}_i \tau} \sigma_i dW_i^P + \sum_i e^{-\overline{\kappa}_i \tau} \left(\kappa_i \mu_i^Y + (\overline{\kappa_i} - \kappa_i) Y_{it}\right) dt \\
&= r dt + \sum_i e^{-\overline{\kappa}_i \tau} \sigma_i dW_i^P + \sum_i e^{-\overline{\kappa}_i \tau} \left(\overline{\kappa_i} \mu_i^Y + (\overline{\kappa_i} - \kappa_i)(Y_{it} - \mu_i^Y)\right) dt \ .
\end{aligned}$$

We conclude that the expected value of the drift for a (long) rolling futures strategy under the stationary distribution is:

$$r + E\Big[\sum_i e^{-\overline{\kappa}_i \tau} \left(\overline{\kappa_i} \mu_i^Y + (\overline{\kappa_i} - \kappa_i)(Y_{it} - \mu_i^Y)\right)\Big] = r + \sum_i e^{-\overline{\kappa}_i \tau} \overline{\kappa_i} (\mu_i - \bar{\mu}_i) \ .$$

Setting

$$m(\tau) = \sum_i e^{-\overline{\kappa}_i \tau} \overline{\kappa_i} (\mu_i - \bar{\mu}_i) \ ,$$

the conclusion is that shorting the long futures ETN has positive drift if and only if

$$m(\tau) < 0 \ .$$

Setting

$$s^2(\tau) = \sum_{ij} \sigma_i \sigma_j \rho_{ij} e^{-(\overline{\kappa}_i + \overline{\kappa}_j)\tau} \ ,$$

we conclude that the Sharpe ratio of the short futures rolling strategy with constant maturity τ is

$$\begin{aligned} SR &= -\frac{\sum_i e^{-\overline{\kappa}_i \tau} \overline{\kappa_i} \mu_i}{s(\tau)} \\ &= \frac{1}{s(\tau)} \frac{\partial ln V^{\tau}_{eq.}}{\partial \tau} - \frac{s(\tau)}{2} \ . \end{aligned}$$

In the next section we estimate the model and provide various numerical results regarding the profitability of rolling futures strategies based on this model.

5 Parameter Estimation

Based on the PCA from Section 2.1 and [Alexander & Korovalis(2013)], we consider the parametric model of Section 4 with $d = 2$. We estimate parameters for this model by fitting to historical futures data. The OU factors can be written as

$$\begin{aligned} X_{1t} &= \mu_1 + \sigma_1 \int_{-\infty}^{t} e^{-\kappa_1 (t-s)} dW^P_{1s} \ , \\ X_{2t} &= \mu_2 + \sigma_2 \int_{-\infty}^{t} e^{-\kappa_2 (t-s)} dW^P_{2s} \ , \end{aligned}$$

where $dW_{1t}^P dW_{2t}^P = \rho dt$. The process $(X_{1t}e^{-\kappa_1\tau}, X_{2t}e^{-\kappa_2\tau})$ is a stationary bivariate OU with mean $(\mu_1 e^{-\kappa_1\tau}, \mu_2 e^{-\kappa_2\tau})$ and covariance matrix

$$cov(X_{1t}e^{-\kappa_1\tau}, X_{2t}e^{-\kappa_2\tau}) \to \begin{pmatrix} \frac{\sigma_1^2 e^{-2\kappa_1\tau}}{2\kappa_1} & \frac{\rho\sigma_1\ \sigma_2 e^{-(\kappa_1+\kappa_2)\tau}}{\kappa_1+\kappa_2} \\ \frac{\rho\sigma_1\sigma_2 e^{-(\kappa_1+\kappa_2)\tau}}{\kappa_1+\kappa_2} & \frac{\sigma_2^2 e^{-2\kappa_2\tau}}{2\kappa_2} \end{pmatrix} ,$$

as $t \to \infty$. Stationarity of the factors make it possible for a likelihood function with enough data to return good parameter estimates.

5.1 Filtering and parameter estimation

Using daily VIX and VX contracts, we estimated the vector of parameters for the bivariate OU model by a least squares fit to the CMF curve. We employed a Kalman filtering approach with parameter estimates being improved iteratively. Denote discrete daily times as

$$t_\ell = t_0 + \ell\Delta t \qquad \text{for } \ell = 0, 1, 2, \ldots, N,$$

where

$$\Delta t = \frac{1}{252} .$$

The input data are futures prices observed daily and CMFs are computed for each day,

$$\mathbf{V}_\ell = \begin{pmatrix} ln(VIX_{t_\ell}) \\ ln(V_{t_\ell}^{\tau_1}) \\ \vdots \\ ln(V_{t_\ell}^{\tau_7}) \end{pmatrix}$$

with $\tau_j = j \times 30$ days for $j = 0, 1, \ldots, 7$. Let the parameters of the OU process by denoted by Θ,

$$\Theta = (\mu_1, \mu_2, \kappa_1, \kappa_2, \bar{\mu}_1, \bar{\mu}_2, \bar{\kappa}_1, \bar{\kappa}_2, \sigma_1, \sigma_2, \rho) .$$

Using the model

$$\mathbf{V}_\ell = H^\Theta X_{t_\ell} + G^\Theta + \varepsilon_\ell \ , \tag{12}$$

$$X_{t_{\ell+1}} = A^\Theta X_{t_\ell} + \mu^\Theta + \Delta W_{\ell+1} \ , \tag{13}$$

the coefficients are

$$H^\Theta = \begin{pmatrix} 1 & 1 \\ e^{-\bar{\kappa}_1\tau_1} & e^{-\bar{\kappa}_2\tau_1} \\ e^{-\bar{\kappa}_1\tau_2} & e^{-\bar{\kappa}_2\tau_2} \\ \vdots & \vdots \\ e^{-\bar{\kappa}_1\tau_7} & e^{-\bar{\kappa}_2\tau_7} \end{pmatrix}$$

$$A^\Theta = \begin{pmatrix} 1-\kappa_1\Delta t & 0 \\ 0 & 1-\kappa_2\Delta t \end{pmatrix} \qquad \mu^\Theta = \begin{pmatrix} \kappa_1\mu_1 \\ \kappa_2\mu_2 \end{pmatrix} \Delta t \ ,$$

and where $G^\Theta = G_1^\Theta + G_2^\Theta$ with

$$G_1^\Theta = ln(V^\infty) - \begin{pmatrix} \bar{\mu}_1 + \bar{\mu}_2 \\ \bar{\mu}_1 e^{-\bar{\kappa}_1\tau_1} + \bar{\mu}_2 e^{-\bar{\kappa}_2\tau_1} \\ \bar{\mu}_1 e^{-\bar{\kappa}_1\tau_2} + \bar{\mu}_2 e^{-\bar{\kappa}_2\tau_2} \\ \vdots \\ \vdots \\ \bar{\mu}_1 e^{-\bar{\kappa}_1\tau_7} + \bar{\mu}_2 e^{-\bar{\kappa}_2\tau_7} \end{pmatrix} \ ,$$

$$G_2^\Theta = -\frac{1}{4}\begin{pmatrix} \frac{\sigma_1^2}{\bar\kappa_1} + \frac{\sigma_2^2}{\bar\kappa_2} + 4\frac{\rho\sigma_1\sigma_2}{\bar\kappa_1+\bar\kappa_2} \\ \frac{\sigma_1^2}{\bar\kappa_1}e^{-2\bar\kappa_1\tau_1} + \frac{\sigma_2^2}{\bar\kappa_2}e^{-2\bar\kappa_1\tau_1} + 4\frac{\rho\sigma_1\sigma_2}{\bar\kappa_1+\bar\kappa_2}e^{-(\bar\kappa_1+\bar\kappa_2)\tau_1} \\ \frac{\sigma_1^2}{\bar\kappa_1}e^{-2\bar\kappa_1\tau_2} + \frac{\sigma_2^2}{\bar\kappa_2}e^{-2\bar\kappa_1\tau_2} + 4\frac{\rho\sigma_1\sigma_2}{\bar\kappa_1+\bar\kappa_2}e^{-(\bar\kappa_1+\bar\kappa_2)\tau_2} \\ \vdots \\ \vdots \\ \frac{\sigma_1^2}{\bar\kappa_1}e^{-2\bar\kappa_1\tau_7} + \frac{\sigma_2^2}{\bar\kappa_2}e^{-2\bar\kappa_1\tau_7} + 4\frac{\rho\sigma_1\sigma_2}{\bar\kappa_1+\bar\kappa_2}e^{-(\bar\kappa_1+\bar\kappa_2)\tau_7} \end{pmatrix} ,$$

and with noise vectors

$$\begin{aligned} \Delta W_\ell^p &\sim iid\, \mathcal{N}\Big(0, Q^\Theta\Big) \\ Q^\Theta &= \Delta t \begin{pmatrix} \sigma_1^2 & \rho\sigma_1\sigma_2 \\ \rho\sigma_1\sigma_2 & \sigma_2^2 \end{pmatrix} \\ \varepsilon_\ell &\sim iid\, \mathcal{N}(0, R) \ , \end{aligned}$$

where $E\varepsilon_{\ell i}\varepsilon_{\ell j} = R_{ij}$.

Given the history $\mathbf{V}_{0:\ell} = (\mathbf{V}_0, \mathbf{V}_1, \ldots, \mathbf{V}_\ell)$, the posterior distribution of X_{t_ℓ} is normal with mean and covariance denoted by

$$\begin{aligned} \widehat{X}_\ell^\Theta &= E^\Theta[X_{t_\ell}|\mathbf{V}_{0:\ell}] \ , \\ \Omega^\Theta &= E^\Theta(X_{t_\ell} - \widehat{X}_\ell^\Theta)(X_{t_\ell} - \widehat{X}_\ell^\Theta)^{\text{tr}} \ , \end{aligned}$$

and which are given by the Kalman filter,

$$\widehat{X}_{\ell+1}^\Theta = A^\Theta \widehat{X}_\ell^\Theta + \mu^\Theta + K^\Theta\Big(\mathbf{V}_{\ell+1} - H^\Theta(A^\Theta \widehat{X}_{t_\ell}^\Theta + \mu^\Theta) - G^\Theta\Big) \ ,$$

where

$$
\begin{aligned}
\Omega^{\Theta} &= (I - K^{\Theta} H^{\Theta}) \widetilde{\Omega}^{\Theta} \\
K^{\Theta} &= \widetilde{\Omega}^{\Theta} (H^{\Theta})^{\mathrm{tr}} \left(H^{\Theta} \widetilde{\Omega}^{\Theta} (H^{\Theta})^{\mathrm{tr}} + R \right)^{-1} \\
\widetilde{\Omega}^{\Theta} &= A^{\Theta} \widetilde{\Omega}^{\Theta} (A^{\Theta})^{\mathrm{tr}} \\
&\qquad - A^{\Theta} \widetilde{\Omega}^{\Theta} (H^{\Theta})^{\mathrm{tr}} \left(H^{\Theta} \widetilde{\Omega}^{\Theta} (H^{\Theta})^{\mathrm{tr}} + R \right)^{-1} H^{\Theta} \widetilde{\Omega}^{\Theta} (A^{\Theta})^{\mathrm{tr}} + Q^{\Theta} \ .
\end{aligned}
$$

We denote the innovations process as

$$
\nu_{\ell}^{\Theta} = \mathbf{V}_{\ell} - H^{\Theta} (A^{\Theta} \widehat{X}_{\ell-1}^{\Theta} + \mu^{\Theta}) - G^{\Theta} \ ,
$$

which is an iid normal random variable under the null hypothesis that Θ is the true parameter value,

$$
\nu_{\ell}^{\Theta} \sim iid \, \mathcal{N}\Big(0, H^{\Theta} A^{\Theta} \Omega^{\Theta} (H^{\Theta} A^{\Theta})^{\mathrm{tr}} + H^{\Theta} Q^{\Theta} (H^{\Theta})^{\mathrm{tr}} + R \Big) \ .
$$

Hence there is the log-likelihood function,

$$
\begin{aligned}
& L(\mathbf{V}_{0:N} | \Theta, R) \\
&= -\frac{1}{2} \sum_{\ell=1}^{N} \left\| \left(H^{\Theta} A^{\Theta} \Omega^{\Theta} (H^{\Theta} A^{\Theta})^{\mathrm{tr}} + H^{\Theta} Q^{\Theta} (H^{\Theta})^{\mathrm{tr}} + R \right)^{-1/2} \nu_{\ell}^{\Theta} \right\|^2 \\
&\qquad - \frac{N}{2} ln \left| H^{\Theta} A^{\Theta} \Omega^{\Theta} (H^{\Theta} A^{\Theta})^{\mathrm{tr}} + H^{\Theta} Q^{\Theta} (H^{\Theta})^{\mathrm{tr}} + R \right| \ , \qquad (14)
\end{aligned}
$$

where $\left| \ \cdot \ \right|$ denotes matrix determinant. The maximum likelihood estimate (MLE) is

$$
(\hat{\Theta}, \hat{R})_{mle} = \arg\max_{\Theta, R} L(\mathbf{V}_{0:N} | \Theta, R) \ ,
$$

where the maximization must be performed over a suitably constrained space for Θ and R, e.g., κ's greater or equal to zero and R positive (semi)-definite.

5.2 Results

The MLE for Θ is hard to obtain because the parameter space is large and the likelihood is probably not convex; the parameter Θ has 11 inputs and the matrix R adds many more depending on the number of CMFs used. Hence, the best we can do is implement a gradient search algorithm to look for a local maximizer of the objective given by (14), and then do a qualitative analysis of the output. Our *a priori* knowledge of the future curve gives us intuition on the acceptable range for the model parameters, which will be useful when analyzing an estimate of Θ.

Table 2 has Θ estimates obtained using Matlab's 'fmincon' applied to the objective in (14), with input data being the daily closings of VIX and CMFs from February 2011 to December 2016 for the first set of estimates, and July 2007 to July 2016 for the next set of estimates. The table shows how the search's output depends on the input data; the output depends on the range of dates and also the choice of CMFs. Overall, these estimates indicate an equilibrium volatility for the differential of $ln(VIX)$ around 90%, and a very ephemeral and volatile second mode, which sporadically gives rise to partial or full backwardation of the CMFs. These statistical parameters allow some probability for the appearance of most shapes that were observed in the data.

Overfitting can be an issue for models that have many parameters. The MLE obtained by using the Kalman filter does not overfit because of its *a priori* preference for the underlying OU dynamics of equations (12) and (13). An alternative to MLE is the ordinary least-squares estimator,

$$\widehat{X}_\ell^{\Theta,lsq} = \Big((H^\Theta)^{\text{tr}} H^\Theta\Big)^{-1} (H^\Theta)^{\text{tr}} \Big(\mathbf{V}_\ell - G^\Theta\Big) \ ,$$

but this estimator could overfit as there would be no penalization if $(\widehat{X}_\ell^{\Theta,lsq})_{\ell=1,\dots,N}$ did not resemble an OU process. However if the parameter estimation is complete, then $\widehat{X}_\ell^{\Theta,lsq}$ will be a good estimator. Figure 7 compares the post-Θ-estimation time series of $\widehat{X}_\ell^{\Theta,lsq}$ and $\widehat{X}_\ell^{\Theta}$, from which it can be seen that they are very close.

To gain a sense of the quality of fit, Figure 8 shows the residuals of the filtered model. Namely, if the model is correct and the correct parameter values used, then the innovations and the estimated $\widehat{\Delta W}_\ell^{\Theta} = \widehat{X}_{\ell+1}^{\Theta} - A^\Theta \widehat{X}_\ell^{\Theta} - \mu^\Theta$ should have normal distributions centered at zero. Both of these

Estimated Θ
Input data: 2/2011 to 12/2016,
with VIX and CMFs 1M to 7M.

$\bar{\mu}_1$	3.8103	μ_1	3.2957
$\bar{\mu}_2$	-0.7212	μ_2	-0.7588
$\bar{\kappa}_1$	1.1933	κ_1	0.4065
$\bar{\kappa}_2$	10.8757	κ_2	13.1019
σ_1	0.6776		
σ_2	0.8577		
ρ	0.4462		

Estimated Θ
Input data: 2/2011 to 12/2016,
with VIX, 3M and 6M CMFs.

$\bar{\mu}_1$	3.8094	μ_1	3.2524
$\bar{\mu}_2$	-0.7100	μ_2	-0.7155
$\bar{\kappa}_1$	1.0215	κ_1	0.3219
$\bar{\kappa}_2$	10.8739	κ_2	13.0799
σ_1	0.5931		
σ_2	0.9066		
ρ	0.4266		

Estimated Θ
Input data: 7/2007 to 7/2016,
with VIX and CMFs 1M to 7M.

$\bar{\mu}_1$	-6.6216	μ_1	-7.1367
$\bar{\mu}_2$	9.7372	μ_2	9.7608
$\bar{\kappa}_1$	0.6543	κ_1	0.2010
$\bar{\kappa}_2$	5.9052	κ_2	5.9389
σ_1	0.5525		
σ_2	0.9802		
ρ	0.6015		

Estimated Θ
Input data: 7/2007 to 7/2016,
with VIX, 3M and 6M CMFs.

$\bar{\mu}_1$	2.4581	μ_1	1.8685
$\bar{\mu}_2$	0.8002	μ_2	0.7555
$\bar{\kappa}_1$	0.5505	κ_1	0.1081
$\bar{\kappa}_2$	10.0013	κ_2	12.3600
σ_1	0.4294		
σ_2	0.7998		
ρ	0.5073		

Table 2: **Top Panels:** Estimated Θ using daily futures data from February 2011 to December 2016. The terminal value of the log-likelihood was around -10^4. Using VIX and 7 CMFs, the long-term future value is $V^\infty = 25.1206$; using VIX, the 3 month and the 6 month CMF yields $V^\infty = 25.1202$. The parameter search was constrained to match the VIX's statistical mode of 12.64 (i.e. $e^{\mu_1+\mu_2} = 12.64$) and the mean constrained to match the VIX's historical mean of 17.1639 (i.e. $\lim_{t\to\infty} E_0[VIX_t] = exp\left(\mu_1+\mu_2+\frac{1}{4}\left(\frac{\sigma_1^2}{\kappa_1}+\frac{\sigma_2^2}{\kappa_2}+4\frac{\rho\sigma_1\sigma_2}{\kappa_1+\kappa_2}\right)\right) = 17.1639$). The VIX mode is less than both estimates of V^∞, which is consistent with the hypothesis of contango equilibrium.

Bottom Panels: Estimated Θ using daily futures data from July 2007 to July 2016. The terminal value of the log-likelihood was around -10^4. Using VIX and 6 CMFs, the long-term future value is $V^\infty = 27.7316$; using VIX, the 3 month and the 6 month CMF yields $V^\infty = 29.2107$. For this time period the VIX's statistical mode and mean were 13.79 and 21.56, respectively, which were the values used in constraining the parameter search to match these values. Again, the VIX mode is less than both estimates of V^∞, which is consistent with contango equilibrium.

residuals are centered at zero but have heavier tails, which is typical of model fits when dealing with financial data. Nonetheless, the fit is useful going forward.

Finally, Figure 9 shows some CMF curves produce by the estimated parameters, one of which is the equilibrium most-likely curve and the other is a deviation from equilibrium that raises short-term VX prices. Recall the most likely curve from Figure 5 and see that the most-likely curve in Figure 9 is close.

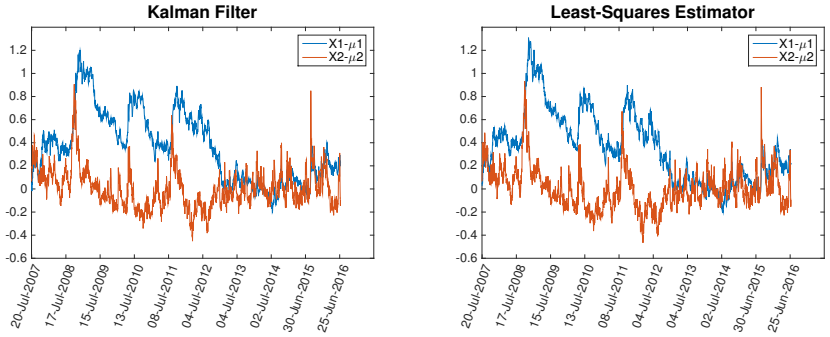

Figure 7: The Kalman filter $\widehat{X}_\ell^\Theta$ and the least-square estimator $\widehat{X}_\ell^{\Theta,lsq}$, both using the estimated Θ from the bottom right panel of Table 2. Unlike the Kalman filter, the least-squares estimator has no *a priori* preference to resemble an OU process, and hence an MLE for Θ based on $\widehat{X}_\ell^{\Theta,lsq}$ may overfit. However, after the Θ estimation is complete the time series of $\widehat{X}_\ell^\Theta$ and $\widehat{X}_\ell^{\Theta,lsq}$ are close because the residual noise is relatively small.

6 Application to ETNs

6.1 Dynamics of ETNs

Let I^θ denote the value of a continuous rolling futures strategy with average maturity θ. For instance, if $\theta = 30/360$, we can view this as a model for the VXX ETN. We assume the parameters of the model are given in Table 2 of Section 5.1. According to the results of Section 3 on rolling futures,

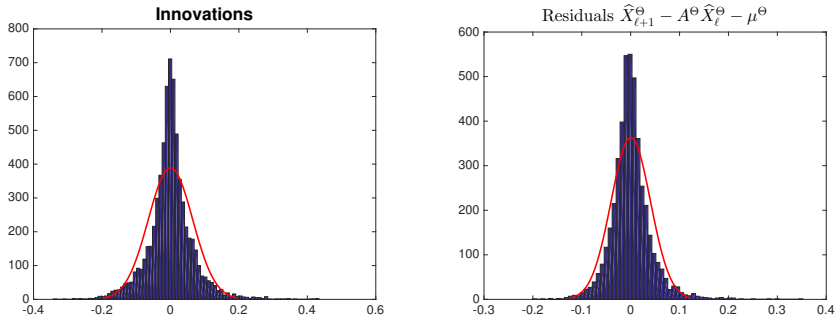

Figure 8: On the left is the histogram of the innovations ν_i using the estimated Θ from the bottom right panel of Table 2, along with a line that is a normal distribution with the same mean and variance. On the right is the histogram of the residuals $\widehat{X}^{\Theta}_{\ell+1} - A^{\Theta}\widehat{X}^{\Theta}_{\ell} - \mu^{\Theta}$ for the same estimate of Θ.

$$\frac{dI^{\theta}}{I^{\theta}} = \sum_{i=1}^{2} \sigma_i e^{-\overline{\kappa}_i \theta} dW_i^P + \sum_{i=1}^{2} e^{-\overline{\kappa}_i \theta} \left(\kappa_i(\mu_i - \bar{\mu}_i) + (\overline{\kappa}_i - \kappa_i)(Y_i - \mu_i^Y) \right) dt \ ,$$

where Y is the re-parameterized OU process that we defined in Section 4.2. This is a Gaussian diffusion process with a stochastic drift. In particular, the equilibrium drift is,

$$m = \sum_{i=1}^{2} e^{-\overline{\kappa}_i \theta} \overline{\kappa}_i (\mu_i - \bar{\mu}_i)$$

and the equilibrium volatility is,

$$s = \sqrt{\sum_{i,j=1}^{2} \rho_{ij} \sigma_i \sigma_j e^{-(\overline{\kappa}_i + \overline{\kappa}_j)\theta}} \ .$$

6.2 Simulated one-month ETN (VXX)

The parameter vectors Θ given in Table 2 produces curves that fit the data and have an equilibrium most-likely curve that is in contango. In this section we take those parameters and simulate various ETNs that roll in VX contracts. It is important for us to remind the reader that the Θ's of Table 2 are estimated *a priori* of any ETN returns, a fact whose significance will be shown in this section where the 2-factor OU model with these parameters will reproduce the historically observed

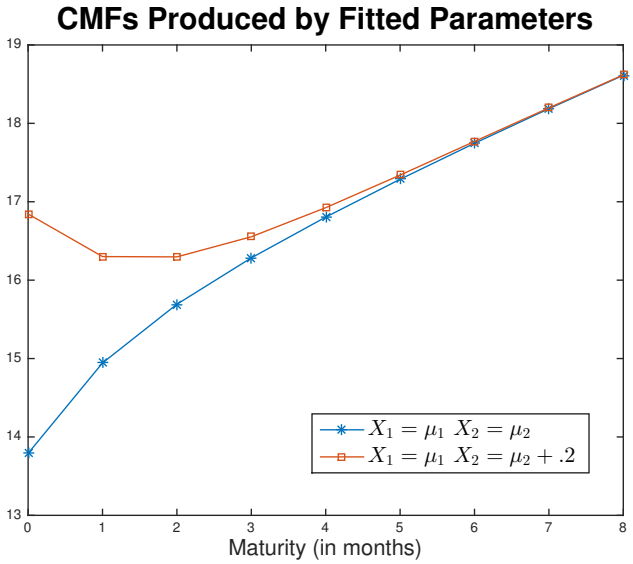

Figure 9: Two possible CMF curves produced by estimated Θ give in the bottom left of Table 2. Taking initial factor values of $X_1 = \mu_1$ and $X_2 = \mu_2$ is the (statistically) most-likely curve. Taking initial factor values $X_2 > \mu_2$ produces a curve with elevated short-term VX contracts.

statistical behavior of the ETNs.

The first ETN that we simulate is a 1 month ETN with a roll yield equal to the slope of the log CMF curve at $\theta = 1$ month. For the 1M long ETN, the estimated Θ from the bottom right panel of Table 2 yield the equilibrium statistics given in Table 3. Figure 10 shows 100 independent sample

1M ETN Equilibrium Statistics

	Jul. 07 to Jul. 16, VIX, CMF 1M to 7M	Jul. 07 to Jul. 16, VIX, CMF 3M, 6M	Feb. 11 to Jul. 16, VIX, CMF 1M to 7M	Feb. 11 to Jul. 16, VIX, CMF 3M, 6M
$\left[\frac{\partial ln V_{eq}^{\tau}}{\partial \tau}\right]_{\tau=\theta}$	0.7389	0.7211	1.0642	0.8473
m	-0.2339	-0.5043	-0.7212	-0.5467
s	1.0050	0.6585	0.8283	0.7753
SR	-0.2328	-0.7658	-0.8706	-0.7051

Table 3: VXX equilibrium statistics given Θ's from Table 2. The estimated model therefore confirms that shorting VXX has positive Sharp ratio between 23% and 87%.

paths from the 2-factor model along with the historical VXX. The VXX appears to lie within the bulk of these sample paths, which is an indication that the model fitted to the CMF curve is able to reproduce historical ETNs. Further analysis of the simulations is seen in the histograms of Figure

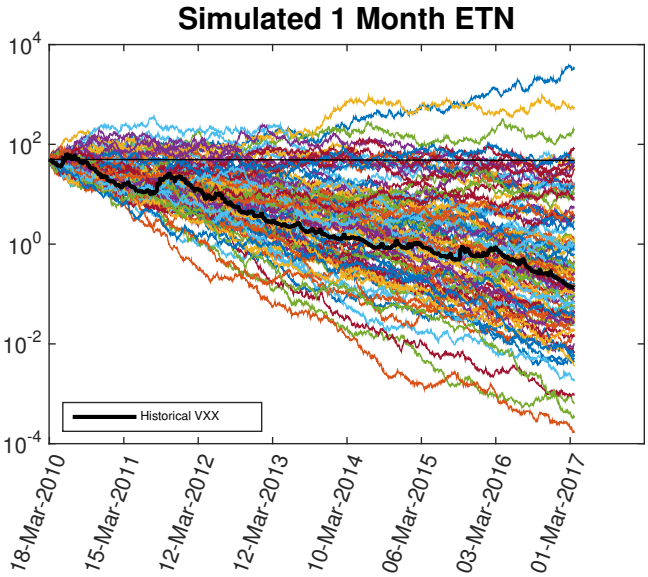

Figure 10: The simulated 1 month ETN along with the historical VXX. The VXX appears to lie within the bulk of these sample paths.

11. These histograms show a normal distribution for the distribution of the estimators for simulated drift, volatility and Sharpe ratios. The equilibrium statistics from Table 3 are clearly at the center of the histograms, and the historically estimated VXX drift, volatility and Sharpe ratio of -0.6418, 0.6301 and -1.0185, respectively, are all within (or close to) the bulk of the simulated estimators. The historical estimators being within their respective histogram's bulk is a statistical non-rejection of the 2-factor model by the VXX time series data.

6.3 Simulated five-month basket ETN (VXZ)

The second ETN that we simulate is a 5 month basket ETN with a roll yield equal to the average slopes of the log CMF curve at $\theta_1 = 4$ months, $\theta_2 = 5$ months and $\theta_3 = 6$ months. The equilibrium drift for this basket ETN is

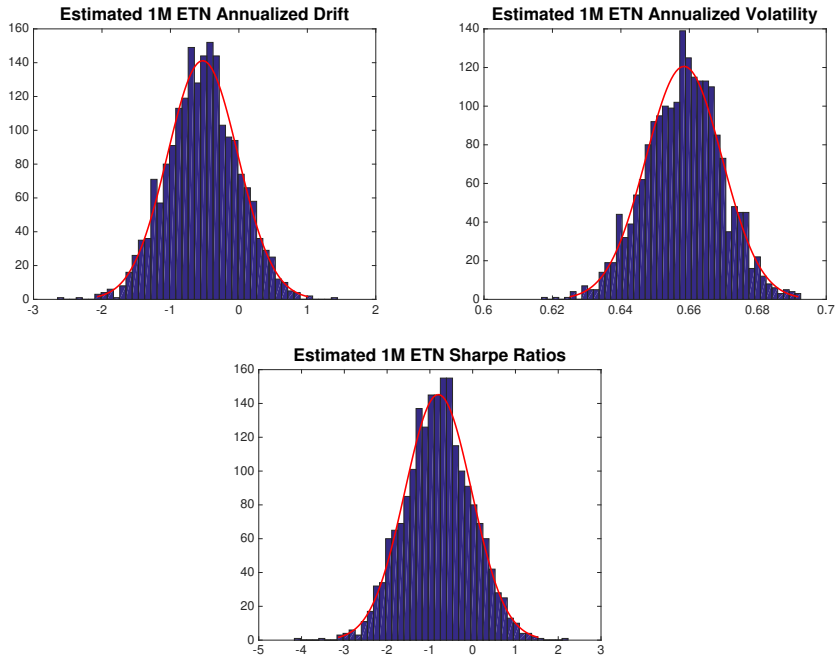

Figure 11: Histograms of VXX estimators obtained using parameters from the bottom right panel of Table 2. **Top Left:** The histogram of estimated drifts for the simulated 1M ETNs. Historically, the 1M ETN VXX recorded an estimated drift of $\hat{m} = \frac{1}{N\Delta t}\sum_\ell \frac{\Delta \text{VXX}_{t_\ell}}{\text{VXX}_{t_\ell}} = -.6418$ between 3/1/2010 and 3/1/2017, which is within the bulk among simulated drift estimators. **Top Right:** For the same time period the historically estimated volatility for the VXX was $\hat{s} = \sqrt{\frac{1}{N\Delta t}\sum_\ell \left(\Delta ln(\text{VXX}_{t_\ell}) - \overline{\Delta ln(\text{VXX})}\right)^2} = 0.6301$, which is at the edge of the bulk among simulated volatility estimators. **Bottom:** The historically estimated VXX Sharpe ratio $\hat{m}/\hat{s} = -1.0185$, which is within the bulk among simulated Sharpe ratio estimates.

$$m = \frac{1}{3}\sum_{j=1}^{3}\sum_{i=1}^{2} e^{-\overline{\kappa}_i\theta_j}\overline{\kappa}_i(\mu_i - \bar{\mu}_i) \ ,$$

but the volatility does not have a closed form so we estimate it empirically with simulation.

For estimated Θ from bottom right panel of Table 2, the equilibrium statistics are given in Table 4. Figure 12 shows 100 independent sample paths from the 2-factor model along with the historical VXZ. The VXZ appears to lie within the bulk of these sample paths, but notice there also appears to be less volatility in VXZ than in the simulated paths. Lower volatility in VXZ is perhaps due

5M Basket ETN Equilibrium Statistics (s and SR from Simulation)

	Jul. 07 to Jul. 16, VIX, CMF 1M to 7M	Jul. 07 to Jul. 16, VIX, CMF 3M, 6M	Feb. 11 to Jul. 16, VIX, CMF 1M to 7M	Feb. 11 to Jul. 16, VIX, CMF 3M, 6M
m	-0.1293	-0.1178	-0.2588	-0.2218
s	0.4829	0.3504	0.4202	0.3951
SR	-0.2677	-0.3362	-0.6158	-0.5613

Table 4: VXZ equilibrium statistics given Θ's from Table 2.

to the fact that it is an average of contracts, and hence there may be some diversification-type effect. Figure 13 shows histograms estimating the distribution of the estimators for simulated drift, volatility and Sharpe ratios. The historically estimated VXZ drift, volatility and Sharpe ratio are -0.2735, 0.3131 and -0.8736, respectively, all of which are within (or close to) the bulk of the simulated estimators, indicating a statistical non-rejection of the 2-factor model by the VXZ time series data.

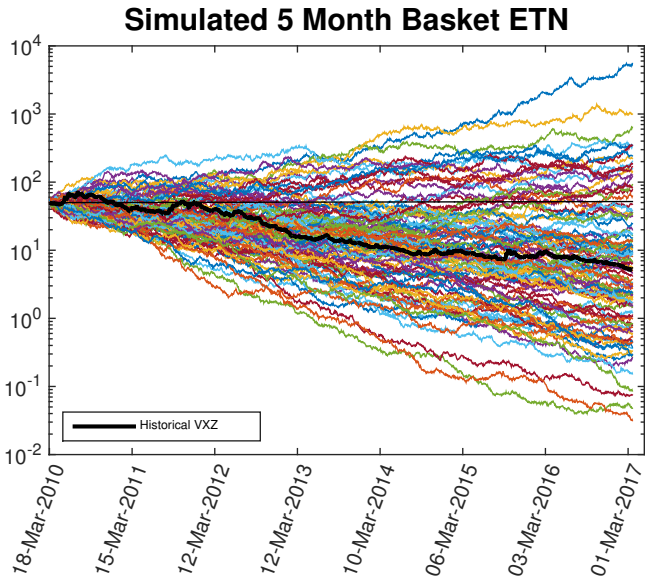

Figure 12: The simulated 5 month basket ETN along with the historical VXZ. The VXZ appears to lie within the bulk of these sample paths.

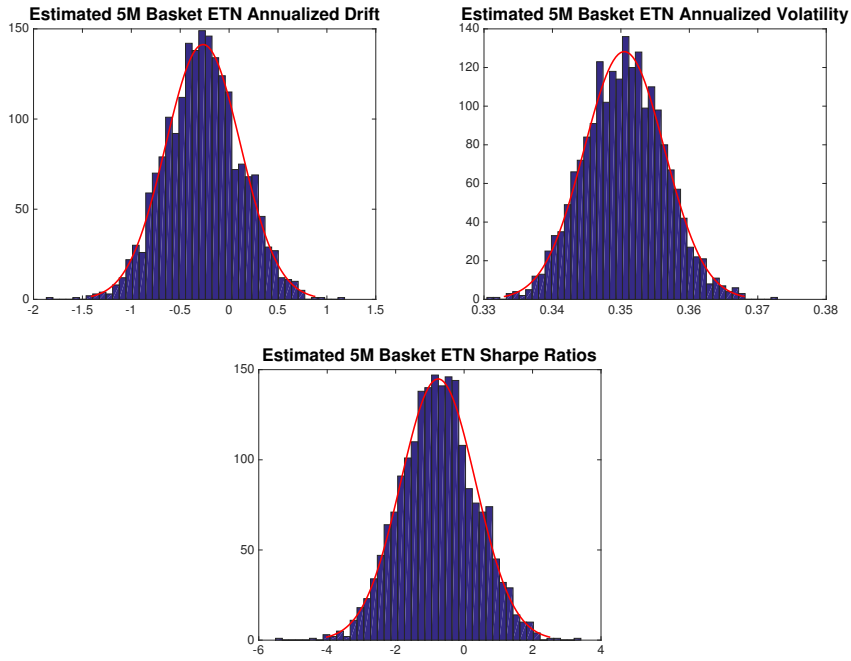

Figure 13: Histograms of VXZ estimators obtained using parameters from the bottom right panel of Table 2. **Top Left:** The histogram of estimated drifts for the simulated 5M basket ETNs with equal weight of 1/3 in the 4,5 and 6 month CMFs. Historically, the 5M basket ETN VXZ recorded an estimated drift of $\hat{m} = \frac{1}{N\Delta t}\sum_\ell \frac{\Delta \text{VXZ}_{t_\ell}}{\text{VXZ}_{t_\ell}} = -0.2735$ between 3/1/2010 and 3/1/2017, which is within the bulk among simulated drift estimators. **Top Right:** For the same time period the historically estimated volatility for the VXZ was $\hat{s} = \sqrt{\frac{1}{N\Delta t}\sum_\ell \left(\Delta ln(\text{VXZ}_{t_\ell}) - \overline{\Delta ln(\text{VXZ})}\right)^2} = 0.3131$, which is near the bulk of simulated volatility estimators. **Bottom:** The historically estimated VXZ Sharpe ratio $\hat{m}/\hat{s} = -0.8736$, which is within the bulk among simulated Sharpe ratio estimates.

6.4 Simulated one-month inverse ETN (XIV)

We now simulate a 1 month inverse ETN with a roll yield equal to the negative slope of the log CMF curve at $\theta = 1$ month. The 1 month inverse ETN has equilibrium statistics that have the opposite sign from those of the 1 month given in Table 3, except for volatility which also has a positive sign. Figure 14 shows 100 independent sample paths from the 2-factor model along with the historical XIV. The XIV appears to lie within the bulk of these sample paths. Figure 15 show histograms estimating the distribution of the estimators for simulated drift, volatility and Sharpe

ratios. The historically estimated XIV drift, volatility and Sharpe ratio are 0.5182, 0.6556 and 0.7904, respectively, all of which are within the bulk of the simulated estimators, and hence there is statistical non-rejection of the 2-factor model by the XIV time series data.

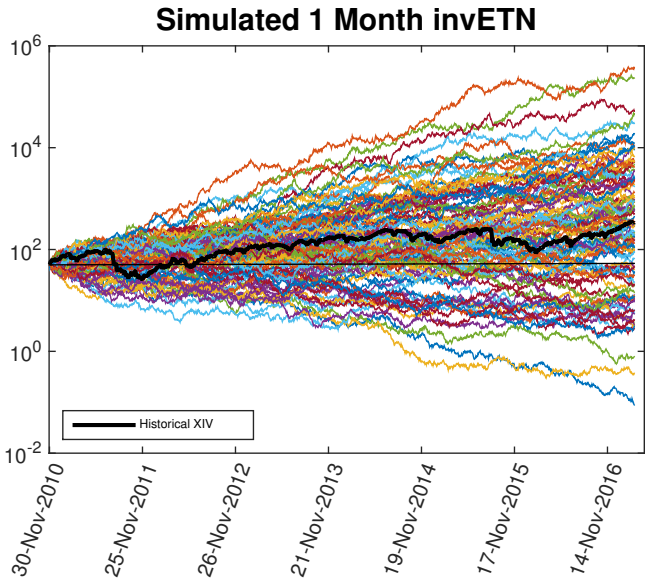

Figure 14: The simulated 1 month inverse ETN along with the historical XIV. The XIV appears to lie within the bulk of these sample paths.

7 Conclusion

We quantified the statistical behavior of VIX constant-maturity-futures in terms of a stochastic model and showed how the time-series of ETP prices depend on the volatility of the CMF curve as and as its slope. Since volatility is often viewed as a stationary, or mean reverting, and the term-structure of CMF is typically in contango, shorting volatility via ETPs should be profitable, in principle. In this paper, we are able to quantify the profitability of selling VIX through ETPs. The results show that although, historically, such strategies have been profitable, the Sharpe ratios achieved are relatively modest (less than the Sharpe ratio for holding the S&P 500). This is due to the spikes in VIX accompanied by CMF backwardation which occur sporadically, but inevitably, in

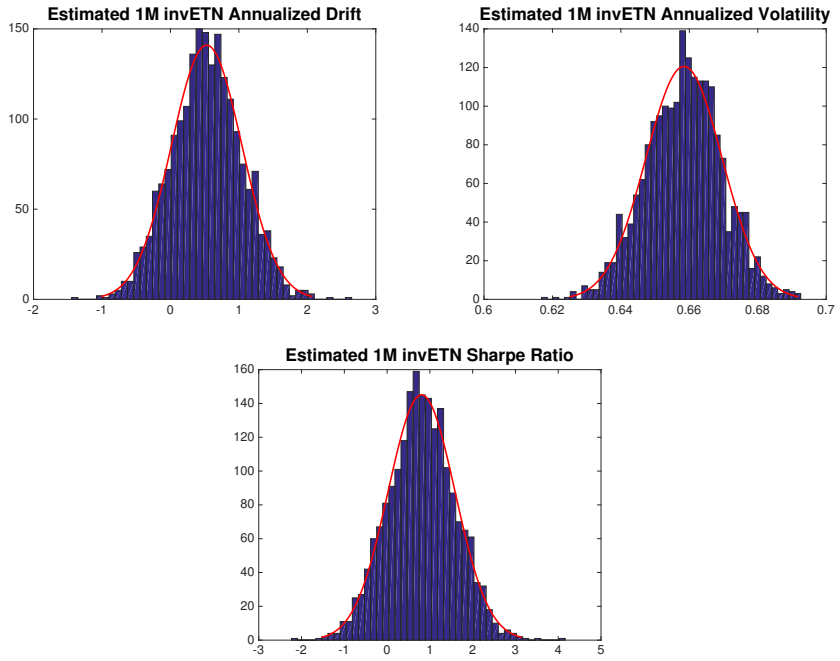

Figure 15: Histograms of estimators obtained using parameters from the bottom right panel of Table 2. **Top Left:** The histogram of estimated drifts for the simulated 1M inverse ETNs. Historically, the 1M inverse ETN XIV recorded an estimated drift of $\hat{m} = \frac{1}{N\Delta t}\sum_\ell \frac{\Delta \text{XIV}_{t_\ell}}{\text{XIV}_{t_\ell}} = 0.5182$ between 11/30/2010 and 3/1/2017, which is within the bulk among simulated drift estimators. **Top Right:** For the same time period the historically estimated volatility for the XIV was $\hat{s} = \sqrt{\frac{1}{N\Delta t}\sum_\ell \left(\Delta ln(\text{XIV}_{t_\ell}) - \overline{\Delta ln(\text{XIV})}\right)^2} = 0.6556$, which is within the bulk among simulated volatility estimators. **Bottom:** The historically estimated XIV Sharpe ratio $\hat{m}/\hat{s} = 0.7904$, which is within the bulk among simulated Sharpe ratio estimates.

the volatility market.

References

[Alexander & Korovalis(2013)] C. Alexander and D. Korovilas (2013) Volatility exchange-traded notes: curse or cure? *Journal of Alternative Investments*, 16(2):52–70.

[Bennett (2014)] C. Bennett (2014) Trading Volatility: Trading Volatility, Correlation, Term Structure and Skew. CreateSpace Independent Publishing Platform, 99 edition, ISBN-10:

1461108756, ISBN-13: 978-1461108757.

[Bergomi (2005)] L. Bergomi (2005) Smile dynamics II. *Risk*, pages 67–73.

[Bergromi (2008)] L. Bergomi (2008) Smile dynamics III. *Risk*, pages 90–96.

[Bergomi (2016)] L. Bergomi (2016) *Stochastic Volatility Modeling*. A Chapman et Hall book. CRC Press.

[Carr & Wu (2009)] P. Carr and L. Wu (2009) Variance risk premiums. *The Review of Financial Studies*, 22:1311–1341, March 2009.

[CBOE (2014)] The Chicago Board Options Exchange (2014)The CBOE Volatility Index - VIX http://www.cboe.com/micro/vix/vixwhite.pdf

[Cox *et al.* (1985)] J.C. Cox, J.E. Ingersoll, S.A. Ross (1985) A theory of the term structure of interest rates *Econometrica* Vol 53 N. 2, 385-408

[Dobi (2014)] D. Dobi (2014) Modeling Systemic Risk in The Options Market *PhD Thesis, Courant Institute of Mathematical Sciences, New York University*

[ETNs (2016)] VIX ETF and ETNs list (2016) http://www.macroption.com/vix-etf-etn-list/

[Gatheral *et al.* (2013)] C. Bayer, J. Gatheral, and M. Karlsmark (2013) Fast Ninomiya-Victoir calibration of the double-mean-reverting model. *Quantitative Finance*, 13(11):1813–1829.

[Laloux *et al.* (2000)] L. Laloux, P. Cizeau, M. Potters, J.-P. Bouchaud (2000) Random Matrix Theory and Financial Correlations *Mathematical MOdels and Methods in Applied Science*

[Litterman & Scheinkman (1991)] R. Litterman, J. Scheinkman (1991) Common Factors Affecting Bond Returns *THe Journal of Fixed Income* p. 54

[Whaley (1993)] R.E. Whaley (1993) Derivatives on Market Volatility: Hedging Tools Long Overdue, *The Journal of Derivatives* 1 , pp. 7184