Trading Signals in VIX Futures

Avellaneda, Li, Papanicolaou, & Wang

Applied Mathematical Finance (2021)

Janos C. R. Füting

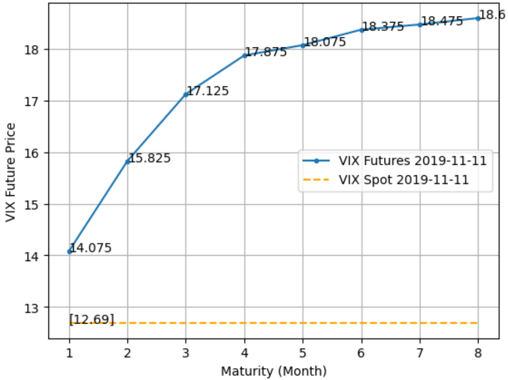
April 3rd, 2024

Outline

VIX Futures and Options

- As we know it is not feasable to replicate the VIX in practice.
- But as people would like to trade it, they came up with a solution: VIX futures (and options)
- Both are cash settled derivatives based on the VIX index as calculated in a ”special settlement auction”
- Maturities range from 1 to 9 months, with less liquid weeklies, and notional of either 1000 or 100 USD per point
- similar markets exist for VSTOXX, VKOSPI, VSMI, etc

Stylised Facts of the VIX Futures Market



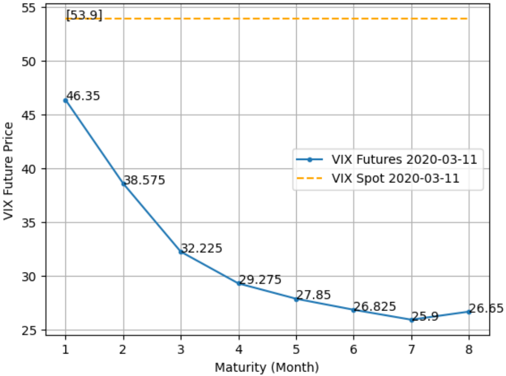


Figure: VIX futures curve on two different days. Left the typical contango structure the market is in most of the time, right the backwardation that is mostly seen when the VIX spikes.

Return Predictability of VIX Futures

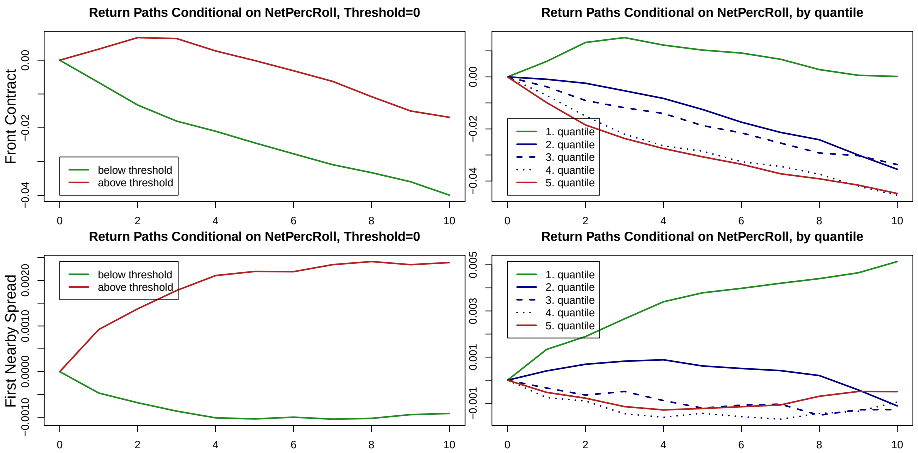


Figure: Cumulative daily returns of front VIX contract/first nearby calendar spread, conditional on signal derived from the futures curve at time $t = 0$.

Return Predictability of VSTOXX Futures

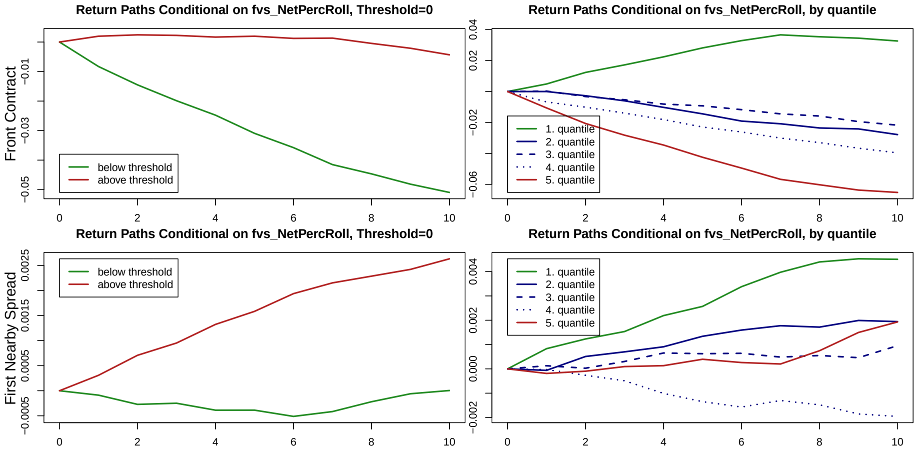


Figure: Cumulative daily returns of front VSTOXX contract/first nearby calendar spread, conditional on signal derived from the futures curve at time $t = 0$.

Paper Introduction

- Empirically, holding VIX futures, or structures products which consist of portfolios of VIX futures, has been an extremely poor choice, with extremely negative returns.
- At the same time, the VIX futures curve is usually in contango, typically deviating only briefly from this state.
- Additionally, it seems that the curve behaves like a Markov process, i.e. all information about its expected shape tomorrow is contained in its shape today.
- The authors want to develop trading strategies based on this Markovian behaviour that exploit the roll (in contango) and mean reversion (in backwardation).
- They choose to do this through trading spreads between portfolios that can be interpreted as constant maturitiy futures.
  - There is no particular reason why one has to approach the problem this way, but is probably mostly motivated by the fact that this is similar to VIX ETPs, and used to be a rather popular trade among bank trading desks.

Data and Processing into Curve Model

- The authors use historical VIX futures curve data between April 2007 and November 2020
- Futures expiration leads to changing statistical properties over their lifetime. To avoid non-stationarity they propose constructing a *constant maturities futures curve*
- The price of the CMF with target maturity θ_i is

$$V_t^i = \omega_t^i F_t^i + (1 - \omega_t^i) F_t^{i+1}$$

with

$$\omega_t^i = \frac{T_{i+1} - t - \theta}{T_{i+1} - T_i}$$

where F_t^i is the actual futures price and T_i is the actual futures expiration date so that $t \leq T_i \leq t + \theta_i \leq T_{i+1}$.

The Markovian Model of the CMF

- It is empirically documented that the VIX futures curve exhibits a (1)-Markovian behaviour, that is the expectd shape of the curve tomorrow seems to only depend on its shape today.
- The authors hence propose a VAR to model the curve, which they parameterise as follows:

$$X_t = \left[\log VIX_t, \log V_t^1, \dots, \log V_t^d, Roll_t^1, \dots, Roll_t^d\right]^T$$

where $Roll_{t+1}^i = \dot{\omega}_t^i \frac{F_{t+1}^{i+1} - F_{t+1}^i}{V_t^i}, \quad \dot{\omega}_t^i = \frac{\omega_{t+1}^i - \omega_t^i}{\delta t}$

$$\Psi_{t+1} = \boldsymbol{\mu} + \boldsymbol{A}\Psi_t + \boldsymbol{Z}_{t+1}$$

where Ψ_t is X_t minus its sample mode.

The Markovian Model of the CMF

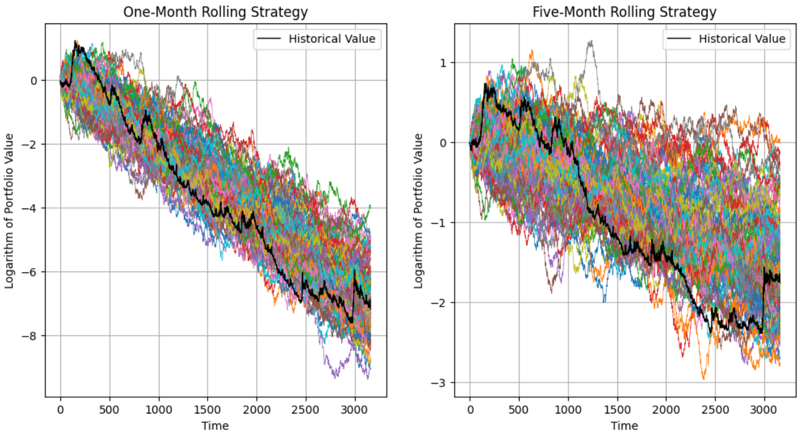


Figure: Historical returns of CMFs versus simulated returns from the fitted VAR, over the study timeframe 2007-04-16 to 2020-11-06. Replication.

Expected returns under the VAR model

- Having fitted the VAR, we can now sample from the stationary distribution $p(X_t)$ and importantly also from the conditional distribution $p(X_{t+1}|X_t)$.
- This we can then use to simulate returns:

$$R_{t+1}(a) = \sum_i a^i \left(Roll^i_{t+1} \Delta t + \frac{V^i_{t+1} - V^i_t}{V^i_t} \right)$$

(recall that X_t is just a bunch of V^i_t and $Roll^i_t$ concatenated).

- And with simulated returns we can calculate expecations of functions of these returns, e.g.

$$U(R) = \max(R, 0) + \gamma \min(R, 0) \qquad U(R) = -\frac{1}{\gamma} \exp(-\gamma R)$$

The Objective Function

The objective function of the NN will be

$$\min_{W,b} \sum_{a \in \mathcal{A}} \frac{1}{N} \sum_{i=1}^{N} \left(Q\left(X_t^{(i)}, a\right) - \frac{1}{M} \sum_{j=1}^{M} U\left(R_{t+1}^{i,j}(a)\right) \right)^2$$

or

$$\min_{W,b} \sum_{a \in \mathcal{A}} \frac{1}{N} \sum_{i=1}^{N} \left(Q\left(X_t^{(i)}, a\right) - U^{-1}\left(\frac{1}{M} \sum_{j=1}^{M} U\left(R_{t+1}^{i,j}(a)\right) \right) \right)^2$$

with $N = 100,000$ and $M = 300$.

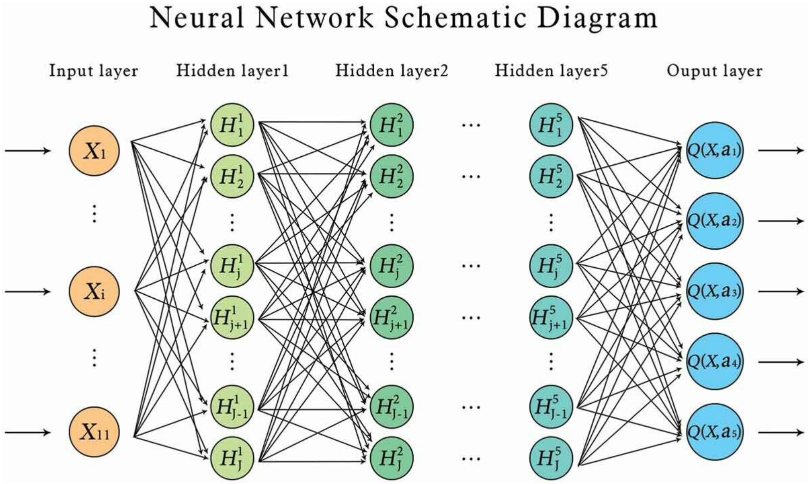
The Objective Function Simplified

Now the previous slide looked mighty fancy, but that is really just

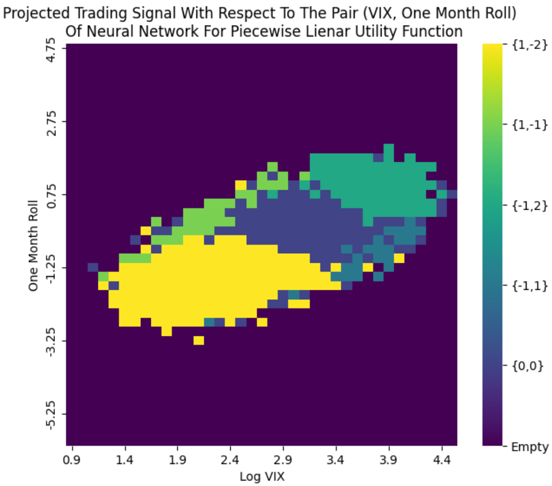
$$MSE(\hat{Y}, \tilde{Y})$$

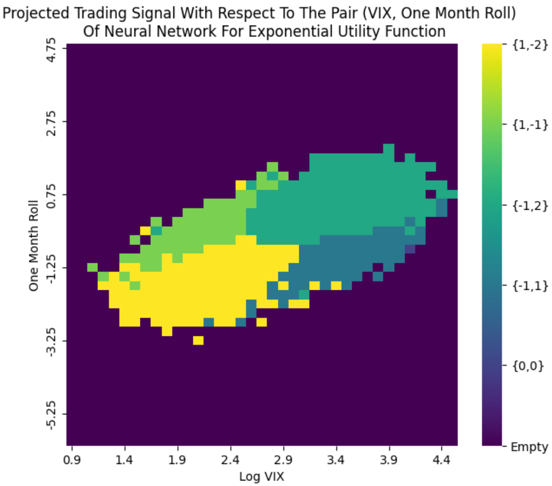
where $\hat{Y}$ is the NN prediction and $\tilde{Y}$ is a Monte Carlo estimate of the utility or certainty equivalent under the VAR distribution...

The Neural Network

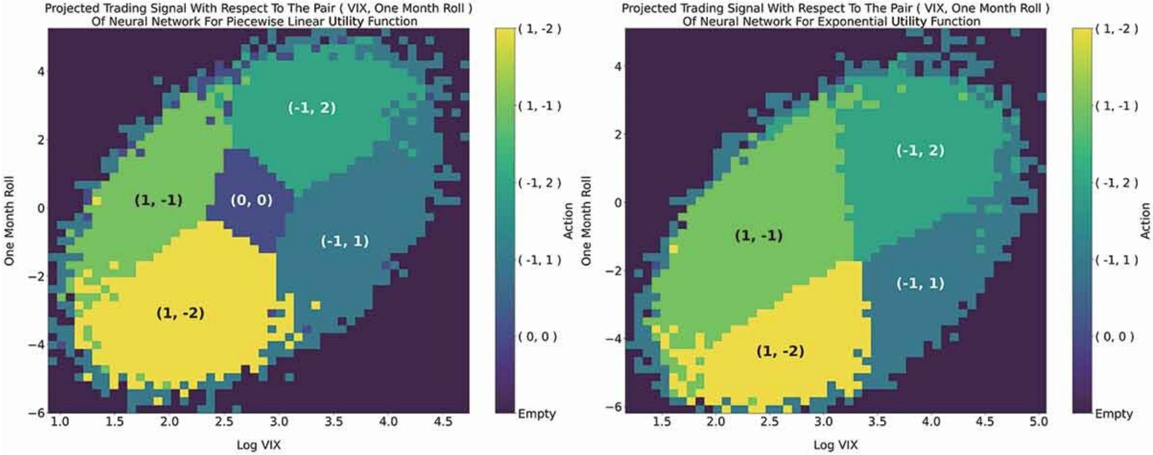


The Trading Strategy - My Result





The Trading Strategy - Their Result



OOS Portfolio Statistics - Their Results

Table 2. Portfolio metrics for out-of-sample tests in k-fold cross validation on trading signal constructed with piece-wise linear utility function (3.5). These metrics are computed from the portfolio returns given by equation (2.10) with no transaction costs.

Fold \ Statistics	$\mathbb{E}[R_t(a(X_t))]$	std $[R_t(a(X_t))]$	Profit $(\%)$	Sharpe Ratio	Maximum Drawdown
0	2.361	0.443	304.264	5.297	−0.196
1	1.195	0.368	145.924	3.215	−0.138
2	4.951	0.447	724.848	11.053	−0.117
3	2.835	0.410	384.387	6.878	−0.214
4	0.854	0.242	108.168	3.474	−0.128
5	1.129	0.361	137.044	3.093	−0.123
6	1.027	0.375	121.582	2.709	−0.156
7	1.415	0.754	130.578	1.862	−0.293
8	0.329	0.302	34.784	1.056	−0.180
9	3.284	0.491	429.191	6.661	−0.240

OOS Portfolio Statistics - My Results

Statistics Fold	$\mathbb{E}[R_t(a(X_t))]$	$std[R_t(a(X_t))]$	Profit %	Sharpe	Max DD
0	0.728	0.343	131.256	2.123	-0.304
1	-0.008	0.254	-4.868	-0.030	-0.191
2	0.463	0.342	66.174	1.352	-0.253
3	0.066	0.497	-6.854	0.132	-0.295
4	-0.227	0.296	-28.755	-0.767	-0.351
5	1.047	0.403	234.690	2.598	-0.222
6	0.053	0.389	-2.460	0.136	-0.351
7	-0.460	0.594	-57.509	-0.774	-0.696
8	0.086	0.297	5.361	0.289	-0.184
9	0.302	0.442	28.733	0.683	-0.561

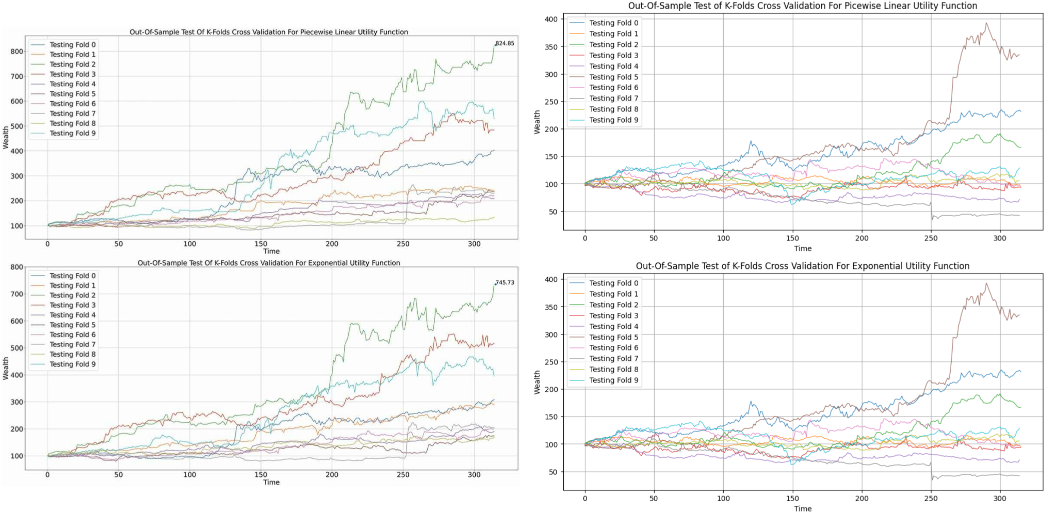


Figure: OOS portfolio values by fold. Left the figure from the paper, right my replication.

Two Views of the Methodology

We can view what the authors are doing from two different angles:

1. Develop a model of the VIX CMF and then solve the optimal allocation decision function though the use of a neural network.
2. Use synthetic data to enable the training of a renforcement learning agent for trading VIX futures.

General Remarks on Replicability

- Multiple figures are wrong.
- The authors mention different timeframes for their results but don't specify which is used where.
- Most information on the training process of the NN are missing.
  - add to this the many other problems with trying to replicate a NN (random seed, software versions, etc)
- Authors don't stick to their own methodology...

Data Quality of VIX Central & Inconsistent Futures Curve Construction

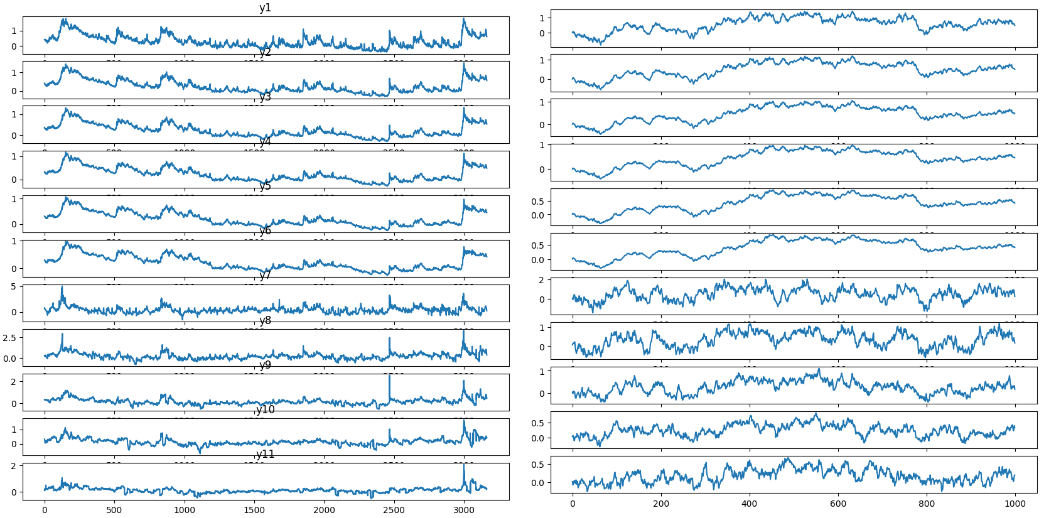
Historical VIX Central data is wrong:

- VIX Central constructs the curve by looking at the expirations at a particular date, and then assigns all prices is knows about to thos expirations starting from the earliest.
- If one of those expirations has no actual price (e.g. did not trade yet), it will get the price of the next expiation, the next expiration will get the one after, and so on ...

The authors throw away all of the nice math they develop for the weights:

- *We take the weights ω_i that appear in Equation (2.2) to be $\omega_i \equiv \omega$ for all i such that there is 100% in the front-month contract as soon as the prior future matures, and then 0% in this front- month at the next maturity date.*
- The effect of this can not be understated, it effectively transforms the data from a constant maturities curve to a random maturities curve...

Failure of Statistical Modelling and Economic Understanding



Failure of Statistical Modelling and Economic Understanding

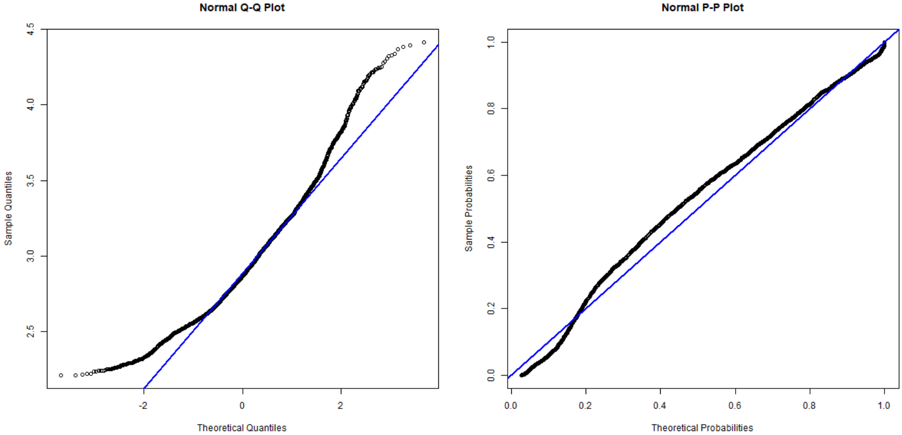


Figure: Empirical VIX distribution versus log-normal model fit.

Incompleteness of the Dynamic Program

- The authors themselves reference Reenforcement Learning literature as the motivation for the neural network approach, but their implementation falls short of the point of RL, to solve dynamic programms. The value function they propose is

$$Q(X_t) = \max_{a \in \mathcal{A}} \mathbb{E}\left[Q(X_t, a)\right]$$

  which is not a dynamic program at all ...
- It is quite easy to make this a proper dynamic program by agumenting the state space by the portfolio one is holding and thus incorporate transaction costs

$$Q(X_t, n_t) = \max_{a \in \mathcal{A}} \mathbb{E}\left[R(X_t, n_t, a) + Q(X_{t+1}, n_{t+1}(a))\right]$$

Thank you for your attention!

References

- Marco Avellaneda, Thomas Nanfeng Li, Andrew Papanicolaou & Gaozhan Wang (2021) Trading Signals in VIX Futures, Applied Mathematical Finance, 28:3, 275-298, DOI: 10.1080/1350486X.2021.2010584