GROUP WORK PROJECT # 1
GROUP NUMBER: 10750

MScFE 692: Capstone Review

FULL LEGAL NAME	LOCATION (COUNTRY)	EMAIL ADDRESS	MARK X FOR ANY NON-CONTRIBUTING MEMBER
Gerardo Jose Suarez Rodriguez	Poland	gerardofingurbe@gmail.com	
Darsh A Patel	USA	darsh@vt.edu	
Giorgio Greto	United Kingdom	giorgiogreto@yahoo.com	

Statement of integrity: By typing the names of all group members in the text boxes below, you confirm that the assignment submitted is original work produced by the group (excluding any non-contributing members identified with an "X" above).	
Team member 1	**Gerardo Jose Suarez Rodriguez**
Team member 2	**Darsh A Patel**
Team member 3	**Giorgio Greto**

Use the box below to explain any attempts to reach out to a non-contributing member. Type (N/A) if all members contributed. **Note:** You may be required to provide proof of your outreach to non-contributing members upon request.
N/A

Group Work Project 1

0.1 Step 1

Extend the PCA monitoring framework in lesson 4.3 with the goal of obtaining a systemic risk indicator combining signals of correlation breakdown and of widening credit spread The computational framework adopted in lesson § 4.3 made use of nine ETFs of the SPDR series, issued by State Street to target specific industry and service sectors in the economy, with each ETF composed exclusively of stocks that are within the S&P500 index. The computations start from January 2007 and extend until the present date. That is why only the nine ETFs already existing in 2007, out of the full suite of eleven SPDR ETFs, were used in the analysis.

Beyond the SPDR ETFs, extra tickers were tracked to follow the volatility of the markets (^VIX), the 10-year Treasury bonds yields (^TNX), and an ETF from iShares reproducing the yield of investment grade corporate bonds (LQD).

All of the data series above are taken with a daily frequency at markets closing time.

Further, we gather daily yield series from the FRED database, for Baa investment grade $\geq$ 20-year corporate bonds as graded by Moody's (BAA) and again the 10-year Treasury bond yields (DGS10).

To summarise:

Ticker	Description
	SPDR ETFs
`XLB`	S&P500 Materials stocks
`XLE`	S&P500 Energy stocks
`XLF`	S&P500 Financial stocks
`XLI`	S&P500 Industrial stocks
`XLK`	S&P500 Technology stocks
`XLP`	S&P500 Consumer Staples stocks
`XLU`	S&P500 Utilities stocks
`XLV`	S&P500 Healthcare stocks
`XLY`	S&P500 Consumer Discretionary stocks
	Other market data
`^VIX`	Market volatility index
`^TNX`	10-year Treasury yields index
`LQD`	Liquid, investment-grade corporate bond yields
	FRED data
`BAA`	Moody's Baa-rated corporate bond yields

Ticker	Description
DGS10	10-year Treasury yields

Below, the shape of the `pandas` DataFrame for the 9 SPDR ETFs downloaded, sporting for dimensions the number of daily observations for the period 2007-2025 (after data cleaning) × the 9 ETFs.

```
Shape of downloaded market data of SPDR ETFs:  (4687, 9)
```

We see that the number of observations for the remaining data series extracted from markets agrees with the previous database at 4684 datapoints:

```
Shape of downloaded market data of volatility and fixed income indices:  (4687,
3)
```

FRED series `BAA`, although nominally taken with daily frequency as declared on the FRED website, in reality exhibits monthly frequency when extracted from the database.

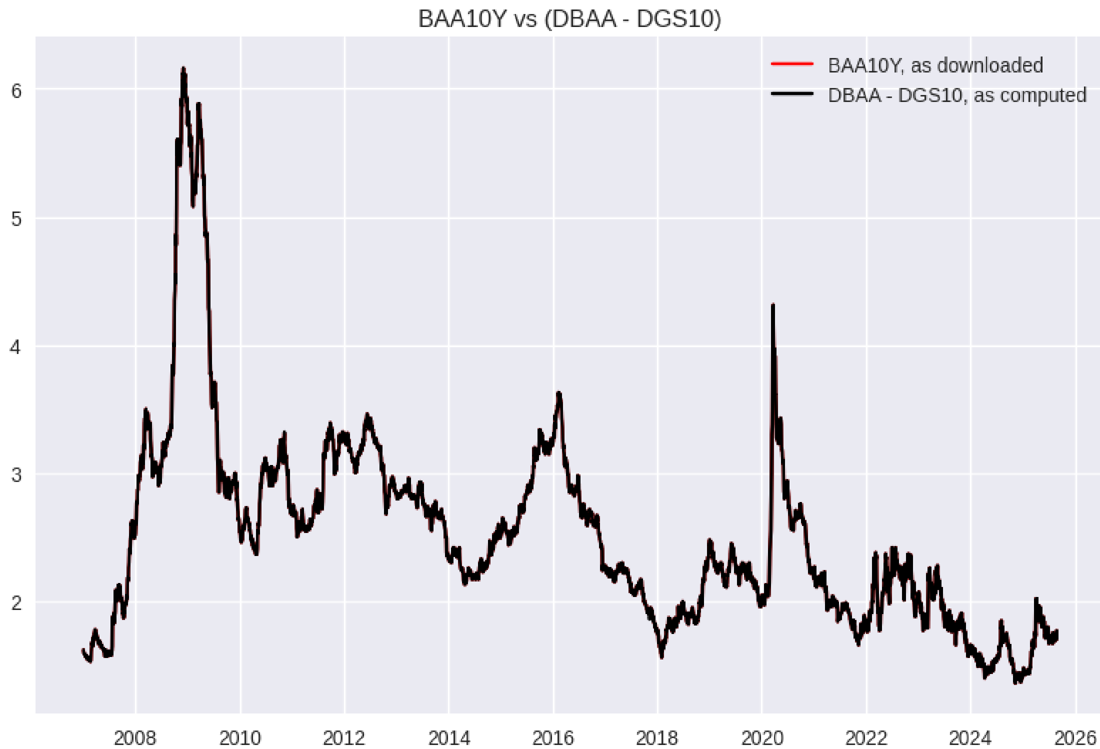
```
id                                                                       BAA
realtime_start                                                    2025-08-21
realtime_end                                                      2025-08-21
title                                  Moody's Seasoned Baa Corporate Bond Yield
observation_start                                                 1919-01-01
observation_end                                                   2025-07-01
frequency                                                            Monthly
frequency_short                                                            M
units                                                                Percent
units_short                                                                %
seasonal_adjustment                                   Not Seasonally Adjusted
seasonal_adjustment_short                                                NSA
last_updated                                          2025-08-01 10:16:05-05
popularity                                                                71
notes                            These instruments are based on bonds with matu…
dtype: object
```

A workaround is found in downloading FRED's `DBAA` series instead, with the initial letter D evidently standing for *daily*.

Given that the Treasury yields data `DGS10` have more observations than all other series downloaded, in fact spanning over days when markets were closed, we will fill the missing datapoints in `DBAA` with the previous most recent observation available.

There also exist a FRED series `BAA10Y` = `DBAA - DGS10` which tracks the credit spread of investment-grade corporate bonds out of the box. The graph below shows `BAA10Y` is obtained exactly as `DBAA - DGS10`.

```
Shape of spread data from Fred:  (4865, 3)
```



For the sake of comparing the credit spread obtained from FRED data (`BAA10Y`) with the one obtained from market data (`LQD - ^TNX`), we first adapt the former to have the same number of observations of the latter.

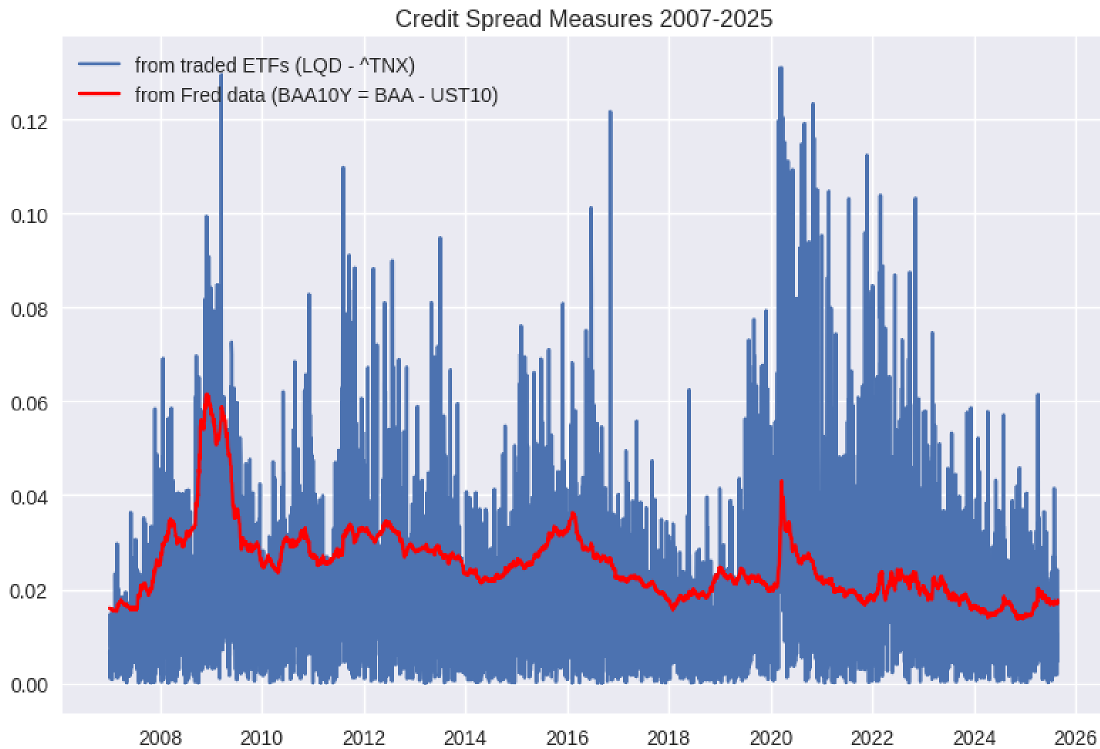
```
shape of spread data from markets:  (4687,)
```

```
shape of spread data from FRED:  (4687,)
```

The credit spreads so obtained by two different sources are plotted in the figure below.

It is evident that the market measure (blue) is much more noisy than that obtained from FRED data (red). We attribute this to the nature of `LQD` as a liquid (highly traded) instrument. Its price is determined by supply and demand dynamics, rather than by the yields of the corporate bonds undergirding it. In turn, this price will reflect the bonds' yields, but only indirectly through market participants' behaviour.

In support of the previous observation, we observe that the absolute value of the noisy market data seems to follow the same patterns as the more regular FRED data. The two measures of credit spread seem to be correlated.



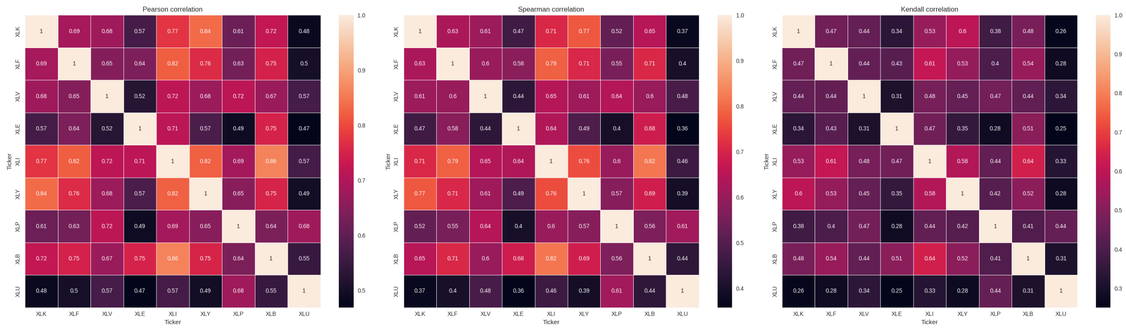
Now, in order to observe the correlations amongst the SPDR ETFs during the selected 2007-25 time window, we compute rolling correlation matrices for time windows of 60 trading days, advancing one day at the time.

The spectral decomposition (or PCA) of these rolling correlation matrices will yield the eigenvalues of the matrices. The largest eigenvalue is found to represent the *market factor* that affects all securities, determining parallel shifts in stock prices.

Its *explained variance ratio* (EVR), i.e. the fraction of total portfolio variance it explains, grows in times of increased market uncertainty. This consists in an undesirable increase in global correlation that could be exploited as a signal for a possibly incipient financial crisis.

We have updated the code in lesson §4.3 to compute the rolling correlation matrices to include not just computing the Pearson (linear) correlation, but also Spearman and Kendall correlations, to track nonlinear effects.

Below, we show the three correlation matrices for the 9 SPDR ETFs, not limited to a few days rolling window, but across the whole time period 2007-25:



We can see from the colorbars in the figures above, showing extreme values, that if nonlinear effects are included (Spearman and Kendall matrices), then the 9 ETFs are overall less correlated than the linear correlation matrix (Pearson) would imply.

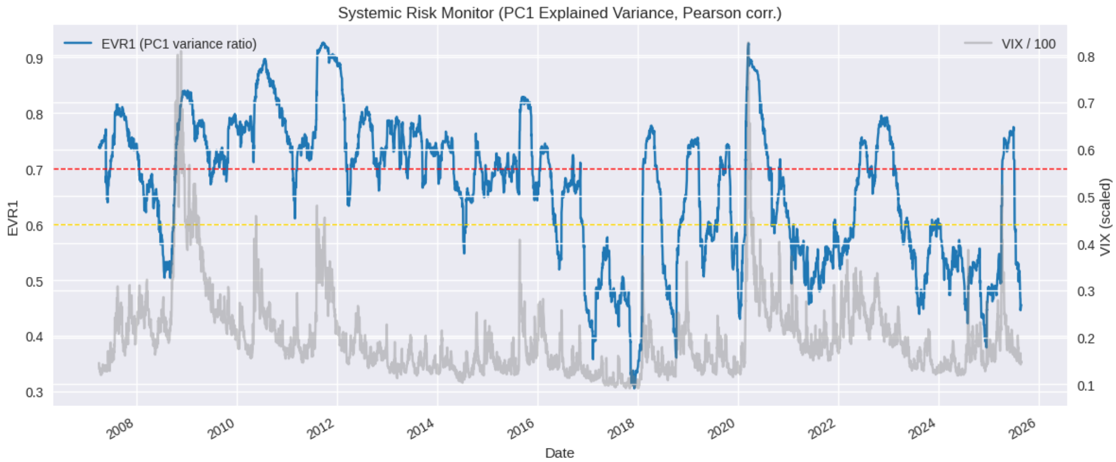
Next, from the rolling 60-days correlations of returns, we compute the EVR for the first eigenvalue (EVR1). We do it for each of the Pearson, Spearman and Kendall correlation matrices.

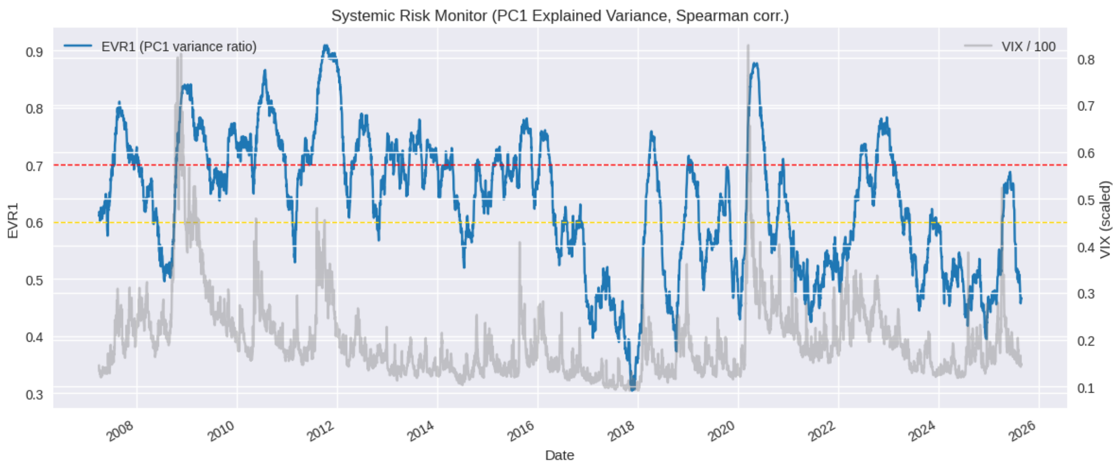
Notice how, on an average home workstation, the computational cost of obtaining the rolling Kendall correlations with Python is 20× more expensive than with the other methods.

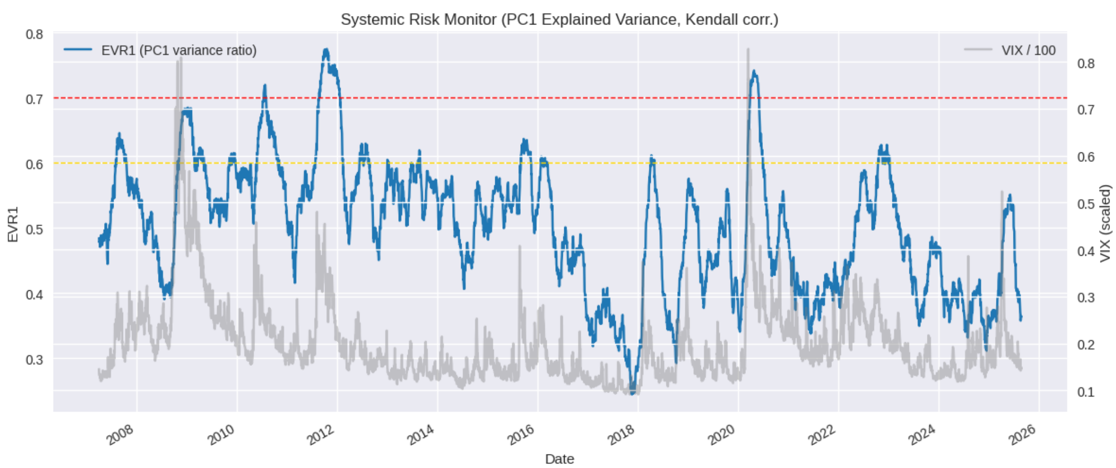
Correlation	Elapsed computation time (s)
Pearson	2.4
Spearman	3.9
Kendall	84.5

From plotting these EVR1s (blue) against time, and in comparison with the ^VIX index (gray), we observe that the EVR1 signal obtained from the Kendall correlation is much more conservative than the previous two. It crosses the highest threshold of correlation (red dotted line) only in three occasions: 2010, 2011-12, and 2020.

It is noticeable that, contrary to the other two correlation methods, the Kendall correlation signal missed the 2007-08 Quant Crash and Great Financial Crisis.





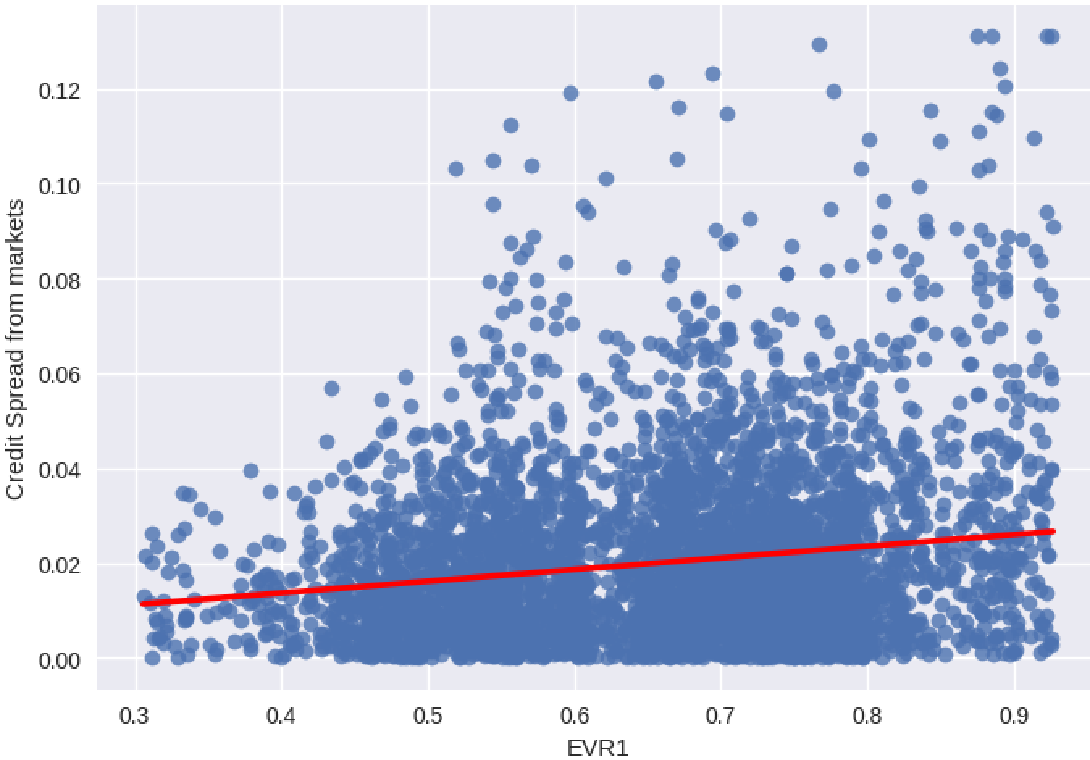


The next step of the analysis consists in checking whether credit spreads and our EVR1 signal are correlated. It would be expected they are as credit spreads widen during systemic crises, just as the proportion of variance explained by the first eigenvalue (EVR1) grows larger.

We have two measures of credit spread, one coming from the FRED database and the other downloaded from the markets (LQD ticker as stand-in for corporate bonds), so we have to select the one better attuned to the EVR1 dataset.

For this purpose, we are going to proceed with a linear regression of the Pearson correlation EVR1 data with both the databases of credit spread at our disposal.

```
                                 OLS Regression Results
================================================================================
======
Dep. Variable:     Credit Spread from markets   R-squared:
0.060
Model:                                    OLS   Adj. R-squared:
0.060
Method:                         Least Squares   F-statistic:
98.88
Date:                        Fri, 29 Aug 2025   Prob (F-statistic):
4.94e-62
Time:                                00:47:34   Log-Likelihood:
12071.
No. Observations:                        4628   AIC:
-2.413e+04
Df Residuals:                            4624   BIC:
-2.411e+04
Df Model:                                   3
Covariance Type:                    nonrobust
==============================================================================
                 coef    std err          t      P>|t|      [0.025      0.975]
------------------------------------------------------------------------------
const         -0.0931      0.008    -11.852      0.000      -0.108      -0.078
EVR1           0.1222      0.008     15.185      0.000       0.106       0.138
EVR2           0.1327      0.013     10.067      0.000       0.107       0.159
EVR3           0.2269      0.022     10.375      0.000       0.184       0.270
==============================================================================
Omnibus:                     1730.064   Durbin-Watson:                   1.628
Prob(Omnibus):                  0.000   Jarque-Bera (JB):             7385.998
Skew:                           1.801   Prob(JB):                         0.00
Kurtosis:                       8.033   Cond. No.                         114.
==============================================================================

Notes:
[1] Standard Errors assume that the covariance matrix of the errors is correctly
specified.

 <AxesSubplot: xlabel='EVR1', ylabel='Credit Spread from markets'>
```



P-values = 0 for the EVR1-2-3 coefficients in the regression are encouraging, however a low value for the R^2 statistic = 0.060 on the contrary signals that a very low proportion of the total variance of the dependent variable, the credit spread, is explained by the EVR coefficients.

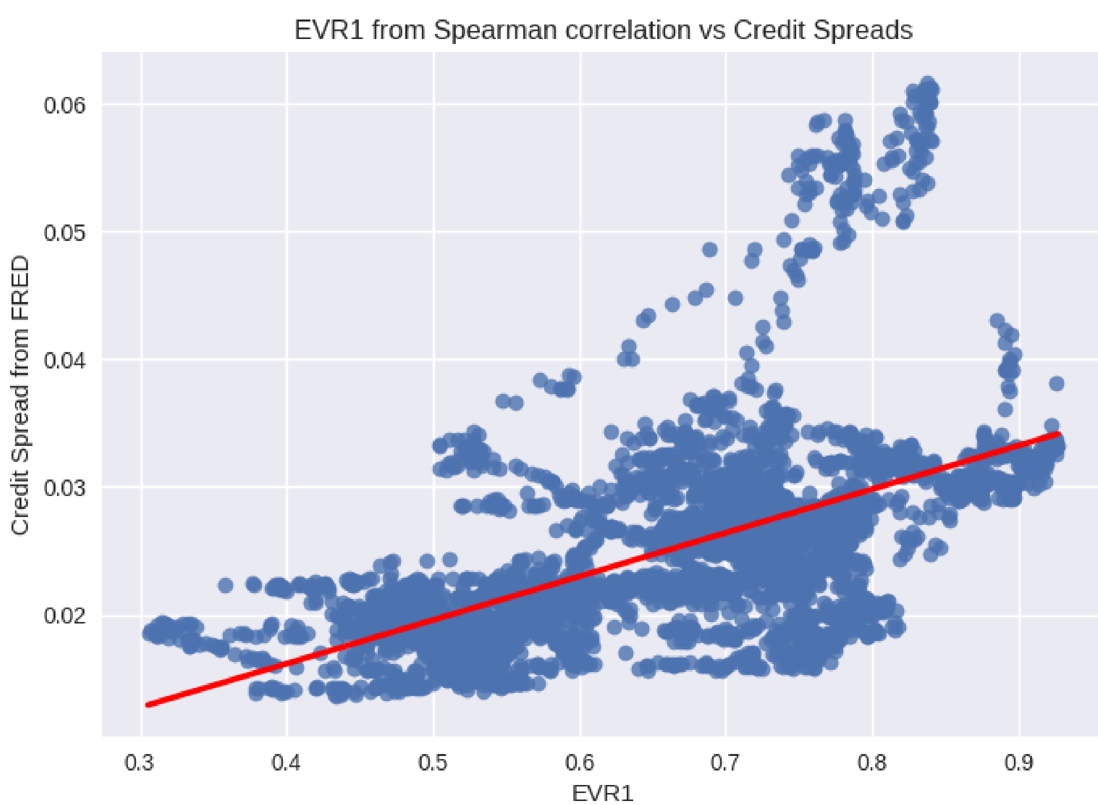
Therefore, we record a mixed outcome for this regression.

Next, much better results are obtained by regressing the EVR measure against the credit spread series from FRED:

```
                            OLS Regression Results
================================================================================
===
Dep. Variable:     Credit Spread from FRED   R-squared:
0.296
Model:                                 OLS   Adj. R-squared:
0.296
Method:                      Least Squares   F-statistic:
648.9
Date:                     Fri, 29 Aug 2025   Prob (F-statistic):
0.00
Time:                             00:47:35   Log-Likelihood:
16657.
No. Observations:                     4628   AIC:
-3.331e+04
```

```
Df Residuals:                    4624   BIC:
-3.328e+04
Df Model:                           3
Covariance Type:            nonrobust
==============================================================================
                 coef    std err          t      P>|t|      [0.025      0.975]
------------------------------------------------------------------------------
const          0.0002      0.003      0.077      0.938      -0.005       0.006
EVR1           0.0364      0.003     12.191      0.000       0.031       0.042
EVR2           0.0093      0.005      1.894      0.058      -0.000       0.019
EVR3          -0.0050      0.008     -0.610      0.542      -0.021       0.011
==============================================================================
Omnibus:                     1888.670   Durbin-Watson:                   0.005
Prob(Omnibus):                  0.000   Jarque-Bera (JB):             9647.938
Skew:                           1.914   Prob(JB):                         0.00
Kurtosis:                       8.948   Cond. No.                         114.
==============================================================================

Notes:
[1] Standard Errors assume that the covariance matrix of the errors is correctly
specified.
```

```
Text(0.5, 1.0, 'EVR1 from Spearman correlation vs Credit Spreads')
```



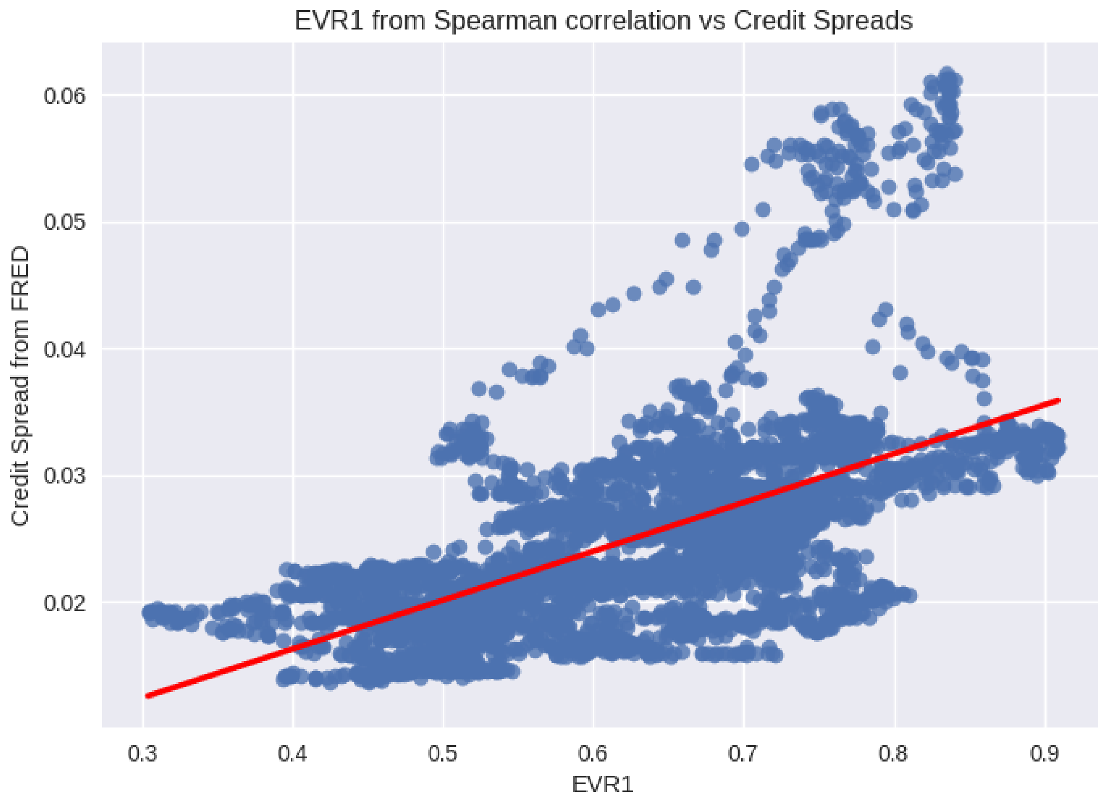
Now the R^2 statistics is more solid, and at the same time the EVR1 coefficient estimation shows a solid p-stat score.

From this information we can infer that credit spreads from Fred are more dependable than their market counterpart, obtained from the LQD ETF.

Proceeding, we complete the picture with the next two regressions which will help selecting the most accurate correlation method for EVR1, in terms of how good it relates to the Fred credit spread.

```
                                 OLS Regression Results
================================================================================
===
Dep. Variable:     Credit Spread from FRED   R-squared:
0.347
Model:                                 OLS   Adj. R-squared:
0.346
Method:                      Least Squares   F-statistic:
818.3
Date:                     Fri, 29 Aug 2025   Prob (F-statistic):
0.00
Time:                             00:47:35   Log-Likelihood:
16829.
No. Observations:                     4628   AIC:
-3.365e+04
Df Residuals:                         4624   BIC:
-3.362e+04
Df Model:                                3
Covariance Type:                 nonrobust
==============================================================================
                 coef    std err          t      P>|t|      [0.025      0.975]
------------------------------------------------------------------------------
const          0.0003      0.002      0.130      0.897      -0.004       0.005
EVR1           0.0390      0.002     16.200      0.000       0.034       0.044
EVR2           0.0233      0.005      5.100      0.000       0.014       0.032
EVR3          -0.0352      0.007     -4.878      0.000      -0.049      -0.021
==============================================================================
Omnibus:                     1735.786   Durbin-Watson:                   0.005
Prob(Omnibus):                  0.000   Jarque-Bera (JB):             8046.635
Skew:                           1.773   Prob(JB):                         0.00
Kurtosis:                       8.400   Cond. No.                         98.5
==============================================================================

Notes:
[1] Standard Errors assume that the covariance matrix of the errors is correctly
specified.
```

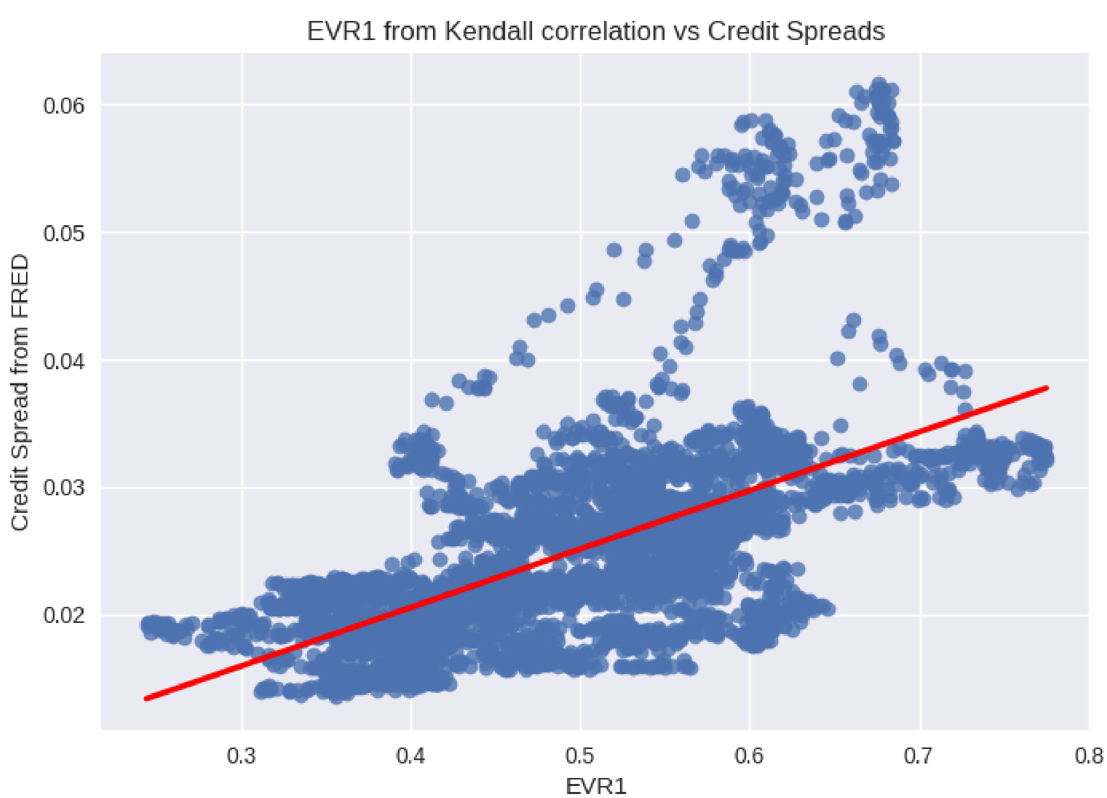
```
Text(0.5, 1.0, 'EVR1 from Spearman correlation vs Credit Spreads')
```




```
                            OLS Regression Results
================================================================================
===
Dep. Variable:     Credit Spread from FRED   R-squared:
0.349
Model:                                 OLS   Adj. R-squared:
0.349
Method:                      Least Squares   F-statistic:
826.2
Date:                     Fri, 29 Aug 2025   Prob (F-statistic):
0.00
Time:                             00:47:36   Log-Likelihood:
16837.
No. Observations:                     4628   AIC:
-3.367e+04
Df Residuals:                         4624   BIC:
-3.364e+04
Df Model:                                3
Covariance Type:                 nonrobust
==============================================================================
```

```
                 coef    std err          t      P>|t|      [0.025      0.975]
------------------------------------------------------------------------------
const       7.413e-05      0.002      0.032      0.975      -0.004       0.005
EVR1           0.0481      0.003     19.101      0.000       0.043       0.053
EVR2           0.0348      0.006      6.026      0.000       0.023       0.046
EVR3          -0.0385      0.009     -4.123      0.000      -0.057      -0.020
==============================================================================
Omnibus:                     1752.466   Durbin-Watson:                   0.005
Prob(Omnibus):                  0.000   Jarque-Bera (JB):             8076.448
Skew:                           1.795   Prob(JB):                         0.00
Kurtosis:                       8.385   Cond. No.                         117.
==============================================================================

Notes:
[1] Standard Errors assume that the covariance matrix of the errors is correctly
specified.
```

```
Text(0.5, 1.0, 'EVR1 from Kendall correlation vs Credit Spreads')
```



All three correlation methods, Pearson, Spearman and Kendall, achieve a p-statistic = 0 for the EVR1 coefficient. Values of the R^2 statistic for Spearman and Kendall are better than Pearson's and more or less equal to each other. The standard error of the EVR1 coefficient in Spearman is slightly lower than Kendall's.

Correlation	EVR1 coefficient	R^2	EVR1 p-stat	EVR1 std err
Pearson	0.0364	0.296	0	0.003
Spearman	0.0389	0.346	0	0.002
Kendall	0.0481	0.349	0	0.003

This, and additionally the previous observation that historical data for EVR1 from Kendall correlation is much more conservative and as such detects less false positives (good) due to noise, but also true positives (bad) such as the 2008 financial crisis, leads us to choose the Spearman correlation as our methodology to extract the EVR1 signal.

Generate a Systemic Risk Indicator (SRI) from correlation breakdown and widening credit spreads An obvious observation that can be made of the above procedure that led us to pick the Spearman correlation matrix to generate the EVR1 metric, is that we applied linear regressions over a correlation measure that tracks nonlinearities.

The analysis would greatly benefit from employing Machine Learning in stead of linear regression to track these nonlinearities, at the price of course of more obfuscation in the selection process.

Indeed, the (Spearman EVR1 - credit spread) linear regression coefficient of 0.0389 is quite feeble, but this also might help justifying the construction of a SRI as linear combination of EVR1 and credit spread. The two terms are expected to correlate more heavily during crises, while a linear combination-based SRI would instead suggest EVR1 and credit spread to be independent from each other. However, a weak correlation when markets are placid might make an acceptable proxy for independence, while the linear combination will still cause the SRI to increase at market phase shifts, just more so than if the rising (EVR1-credit spreads) correlation were to be taken into account.

Therefore, we could define the SRI as following

$$\text{SRI}(\text{credit spread}, \text{EVR1}) = \alpha \cdot \text{credit spread} + \beta \cdot \text{EVR1}$$

where coefficients α, β could be selected from a Least Squares minimisation procedure.

The sum of squared residuals Δ to minimise might be the distance of the SRI from the ^VIX index, as volatility increases during crises, and our goal is to generate an indicator of crises.

$$\Delta = \sum_i \|\text{SRI}_i - \text{VIX}_i\|^2$$

For each observation $i \in [2007, 2025]$ inside the selected timeframe.

We use the `curve_fit` method from `scipy` (based on nonlinear least squares) to calibrate the parameters α and β in SRI. The ^VIX index returns have higher order of magnitude, hence we are going to fit against $\frac{\text{VIX}}{100}$.

```
alpha and beta coefficients:  [6.27575756 0.0663326 ]
```

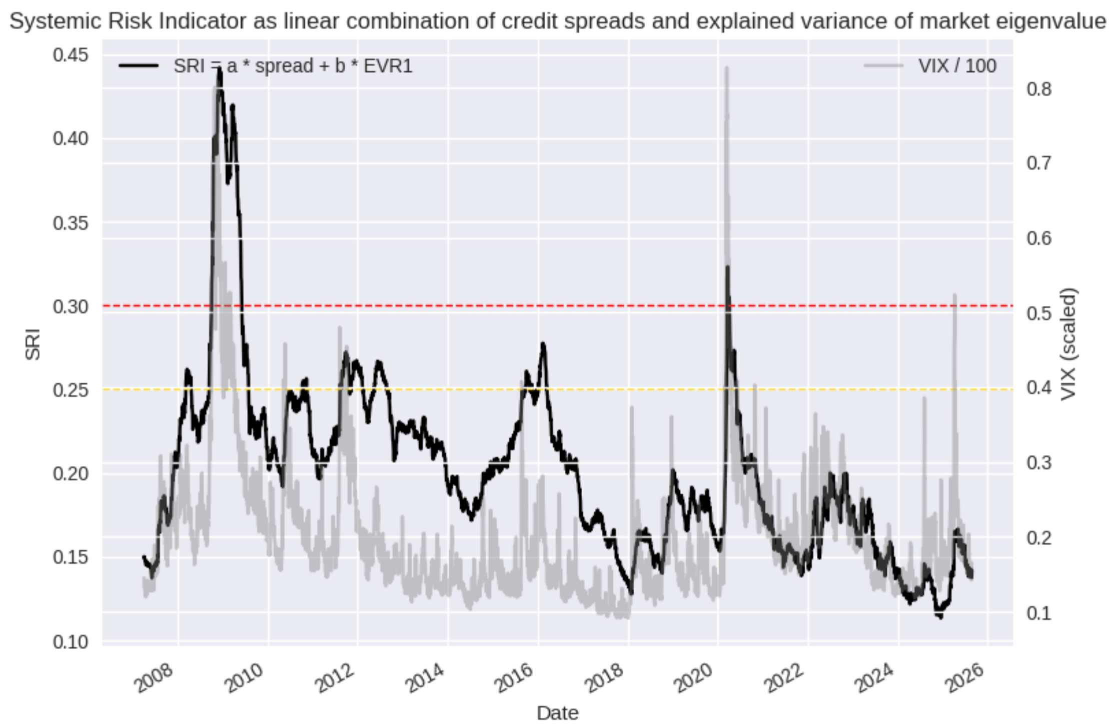
```
stddev of alpha and beta coefficients:  [0.15482733 0.00635288]
```

The procedure yields values of

$$\alpha = 6.276 \pm 0.155$$
$$\beta = 0.066 \pm 0.006$$

This is in line with credit spreads being much smaller in absolute value than the first explained variance of the correlation matrix of returns.

Graphically, the SRI is represented against the ^VIX, below:



Predictably, the SRI mimics well the ^VIX during calm periods.

Outlier values observed during stormy financial periods present a small reaction lag with respect to ^VIX. The financial crisis of 2008 is timely identified by the SRI in the whole of its intensity. Not so much for other periods, like the Covid 2020 crisis, or the increase in US foreign tariff during the first half of 2025. In these cases, the SRI emits tenuous signals when compared to ^VIX. This is why the crisis thresholds were lowered to $0.3 \cdot \text{SRI}$ (red) and $0.25 \cdot \text{SRI}$ (yellow).

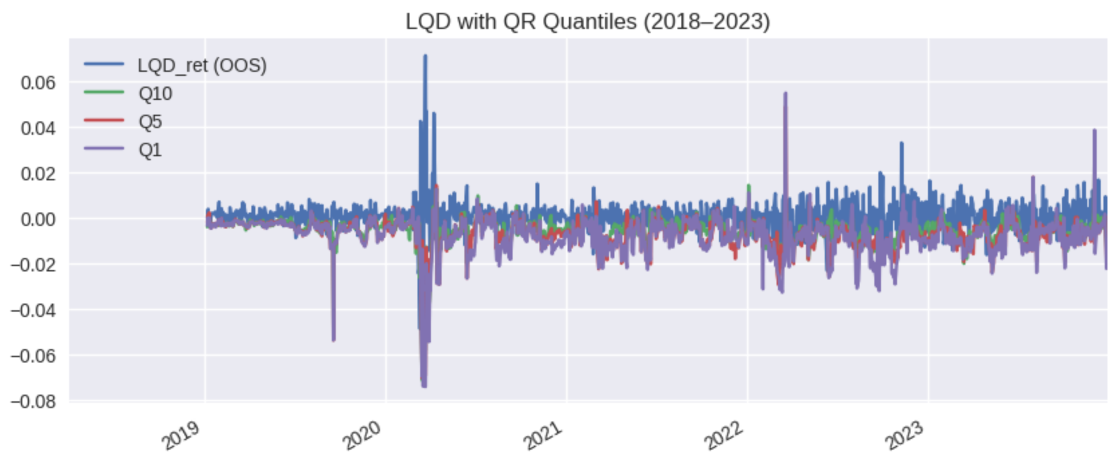
0.1.1 Part B — Quantile Regression with Time-Varying Correlation

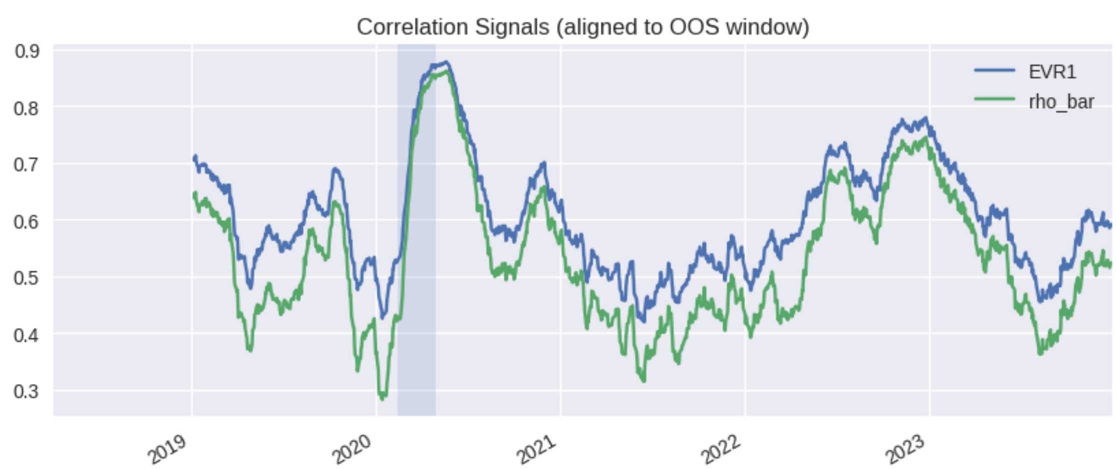
Uses the common loader's variables to fit rolling **Quantile Regression** for $\tau \in \{0.10, 0.05, 0.01\}$ and produce evaluation metrics and figures, all clipped to **2018-04-03 -> 2023-12-31**.

Rolling QR with correlation features adapts, but big regime jumps (like COVID) still cause short-term underestimation. Consider adding a fast-moving volatility proxy (e.g., 21-day realized vol), and include $\frac{\Delta \mathrm{EVR1}}{\Delta \rho}$ to catch rapid correlation shifts.

Higher common-correlation regimes (peaks in 2020, 2023) coincide with periods where left-tail risk is larger, matching the wider/ more negative QR quantiles on the returns chart.

```
[INFO] EVR_WINDOW=63  RHO_WINDOW=63  QR_ROLL_WIN=126
[INFO] EVR1 starts:   2018-07-02 00:00:00 | rho_bar starts: 2018-07-02 00:00:00
[INFO] X starts:      2018-07-05 00:00:00
[INFO] Effective QR window used: 126
[INFO] First OOS pred: 2019-01-07 00:00:00 | OOS points: 1255
tau=0.10  pinball=0.001311  Kupiec(LR,p)=(83.90170871196904,0.0)
Christoffersen(LR,p)=(-21.00243140767132,1.0)
tau=0.05  pinball=0.001060  Kupiec(LR,p)=(190.4456892544049,0.0)
Christoffersen(LR,p)=(-21.657981888211363,1.0)
tau=0.01  pinball=0.000810  Kupiec(LR,p)=(641.9116844342548,0.0)
Christoffersen(LR,p)=(-19.354953346983585,1.0)
```





The quantile envelopes adjust in size according to market regimes, as anticipated. The COVID shock (March–April 2020) results in the most significant negative outcomes—numerous breaches below the Q5 and Q1 curves—demonstrating the model's short-term underestimation of tail risk during an abrupt regime shift. As conditions stabilize through late 2020 and into 2021, the bands constrict, and realized returns primarily fluctuate between Q10 and Q1, indicating improved calibration during calmer periods. In 2022–2023, characterized by rate shocks and liquidity concerns, the band expands once more, leading to occasional exceedances, though significantly fewer than in 2020—this serves as evidence that the rolling QR adjusts, albeit not immediately, to spikes in volatility and correlation.

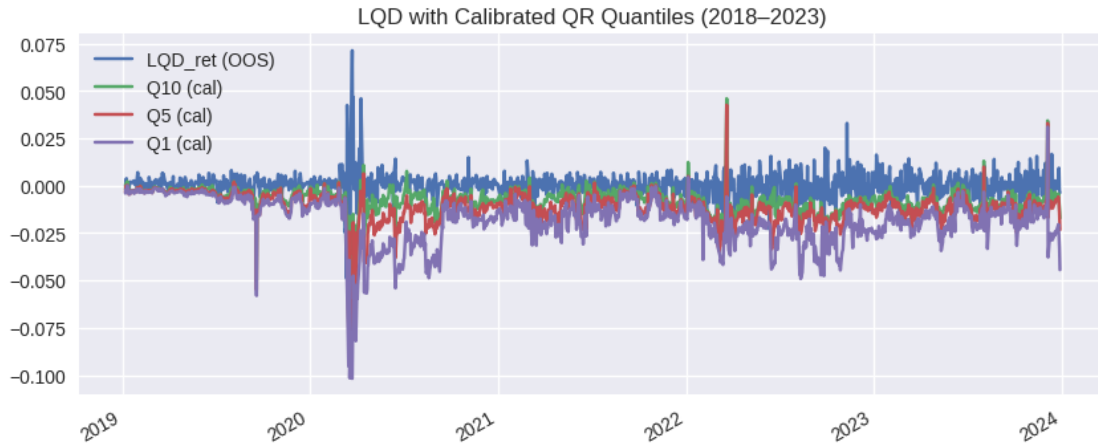
Regarding the levels: the predicted quantiles remain below zero for a substantial portion of the sample, aligning with LQD's sensitivity to left-tail correlation and credit spreads. In instances of local surges (for example, mid-2022 and early 2023), the model preemptively or shortly thereafter reduces the quantile paths, effectively capturing heightened downside risk. For even swifter adaptation during abrupt changes, it may be beneficial to incorporate a short-window realized volatility feature and utilize $\frac{\Delta \mathrm{EVR1}}{\Delta \rho}$ in conjunction with the levels.

EVR1 and rho_bar move in sync, hitting about 0.8+ in early 2020, then going back down before rising again into 2022-2023, and eventually declining. The difference between them (EVR1 being higher than rho_bar) is common: EVR1 looks at how intense the market is ($\frac{\lambda_1}{N}$), which usually is higher than the simple average correlation. The timing matches the returns graph: when the overall correlation goes up, it happens during times when the QR bands drop and spread out (showing more overall movement together, leading to more extreme low values), while a decrease in correlation is linked to tighter bands and less frequent breaks.

```
Q10: hit rate = 18.566% (target 10.0%) | T=1255
Q5: hit rate = 15.458% (target 5.0%) | T=1255
Q1: hit rate = 14.263% (target 1.0%) | T=1255
```

These hit rates mean the model is under-estimating downside risk across the board (predicted lower quantiles are too high / not negative enough). To help with this added a Rolling calibration layer:

```
[CAL] Q10: hit rate = 11.793% (target 10.0%) | T=1255
[CAL] Q5: hit rate = 6.135% (target 5.0%) | T=1255
[CAL] Q1: hit rate = 2.629% (target 1.0%) | T=1255
```



This calibrates each quantile by adding a rolling residual quantile offset so the recent hit rate matches the target.

0.1.2 Student C: extension of a DCC-GARCH framework with funding stress indicators acting as regime change detectors

Dynamical Conditional Correlation generalized autoregressive conditional heteroskedasticity (DCC-GARCH) is a model capable of estimating large time-varying covariance matrices. It does this by using estimating volatilities individually for each asset using GARCH for each separate time series.

The DCC (Dynamic Conditional Correlation) component is what makes the model "integrated" and powerful. It models how the correlations between these assets change over time. In a stable market, the correlation between stocks and bonds might be low or even negative. But during a market crisis, the correlation between almost all assets tends to spike towards 1 which is usually called a correlation spike.

In our case we will model use market sector tracker ETFs and market stress indicators as data for our correlation modeling, the idea is to be able to infer when the probability of correlation spikes are higher ahead of time. The data we are using are:

Sector tracking ETFs:

- XLY: Consumer Discretionary
- XLP: Consumer Staples
- XLE: Energy
- XLF: Financials
- XLV: Health Care
- XLI: Industrials
- XLB: Materials
- XLK: Technology
- XLU: Utilities

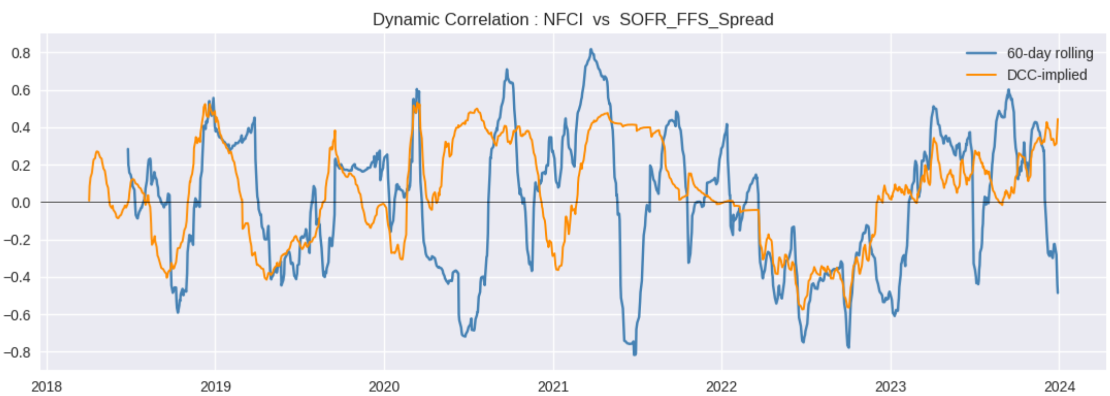
Market stress indicators 1

- TED: Treasury-Eurodollar spread (This one has been discontinued by FRED in favour of Secured Overnight Financing Rate (SOFR))
- SOFR: Secured Overnight Financing Rate
- DFF: Federal Funds Effective Rate
- NFCI: Chicago Fed National Financial Conditions Index
- STLFSI4: St. Louis Fed Financial Stress Index
- VIX:CBOE Volatility Index

The difference between SOFR and DFF can be taken as a proxy of cross currency basis since it indicates dollar liquidity in the repo market. Adding these market indicators, builds on top of the materials provided by WQU.

After obtaining the data we run a $GARCH(1,1)$ model with a Gaussian error term. then we standardise the residuals,and collect the results in a dataframe.

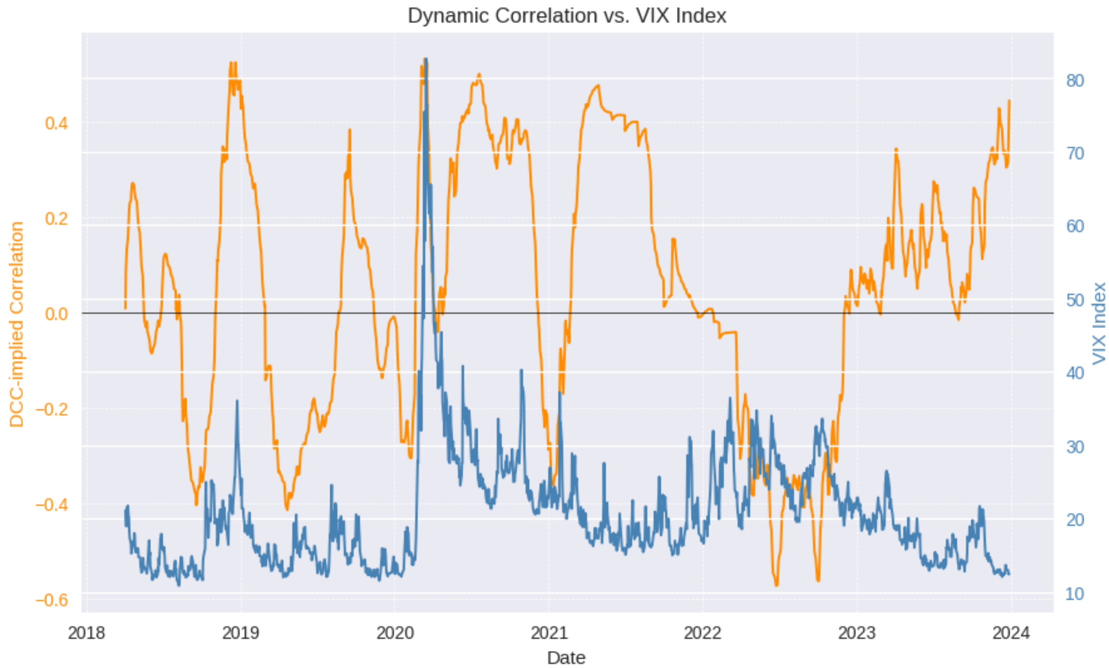
```
DCC alpha=0.050,  beta=0.900,  persistence=0.950
```



A cascade effect in financial markets describes a chain reaction where a shock to one part of the system rapidly spreads to others, creating a self-reinforcing loop of instability. This process often begins with a volatility shock—a sudden, sharp increase in market uncertainty, often triggered by an unexpected event.

A good proxy for this is the VIX which is a measure of the expected volatility of the market as a whole, when it is big the markets are quite volatile which usually means it is moving as a whole and correlations are close to one. We observe our DCC indicator (which measures the correlation between NFCI and our cross currency proxy) seems to be able to predict ahead of time when spikes in the vix will be present, we see that from 2023 the indicator breaks down, this is most likely due to the Fed changing its economic policies and entering into a high interest rate regime which had not been seen in many years 3, other correlation pairs were considered but this is the one we found gave the best predictions for low interest rates environments.

```
Text(0.5, 1.0, 'Dynamic Correlation vs. VIX Index')
```



This shock immediately leads to a correlation spike, as a sudden “risk-off” sentiment causes investors to sell all assets indiscriminately, driving their prices down in unison. Assets that were once uncorrelated or negatively correlated now move together, eliminating the benefits of diversification.

The increased volatility and correlation then trigger a funding freeze. As lenders become hesitant to extend credit due to heightened risk, market liquidity dries up. This forces institutions to sell assets to meet their funding needs, which in turn drives prices down further, causing even more volatility. This completes the feedback loop, as the funding freeze exacerbates the initial volatility shock. This cascade demonstrates how seemingly distinct market dynamics are deeply interconnected, creating a vicious cycle that can lead to a systemic crisis.

0.2 Global Systemic Risk Indicator

By integrating the results of the different generalizations we considered, and expanding upon the work of student A, we will create a Global Systemic Risk Indicator

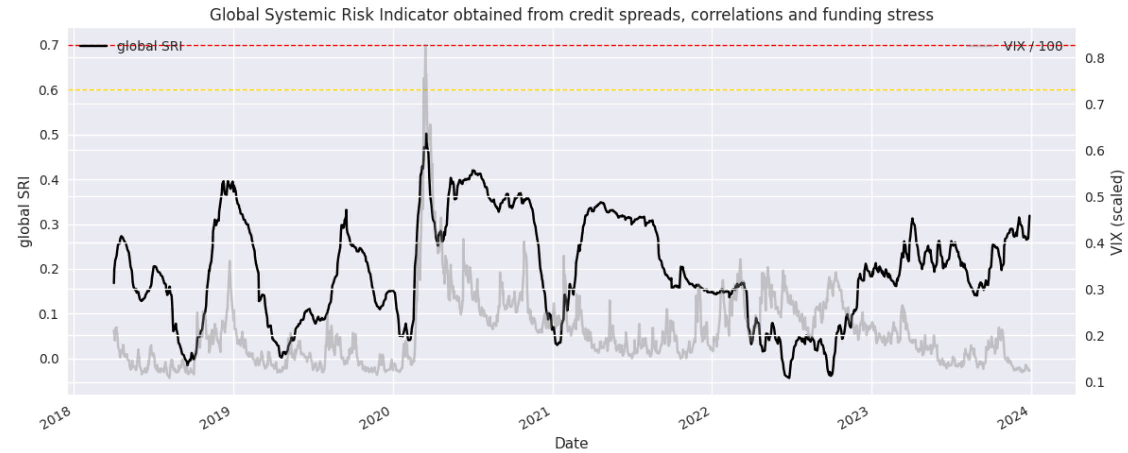
$$\text{GSRI} = \alpha SRI + \beta DCC$$

We will use scipy’s curve fit to produce adequate α, β in such a way that the indicator becomes not only good at predicting possible economic crisis, by going with a simple linear combination of the different indicators we are able to produce an easy to interpret risk indicator, according to Molnar when it comes to algorithms that can have big repercussions, it’s best to have interpretable models so experts can know in what situations they break down.

```
array([1.        , 0.87914484, 0.4       ])
```

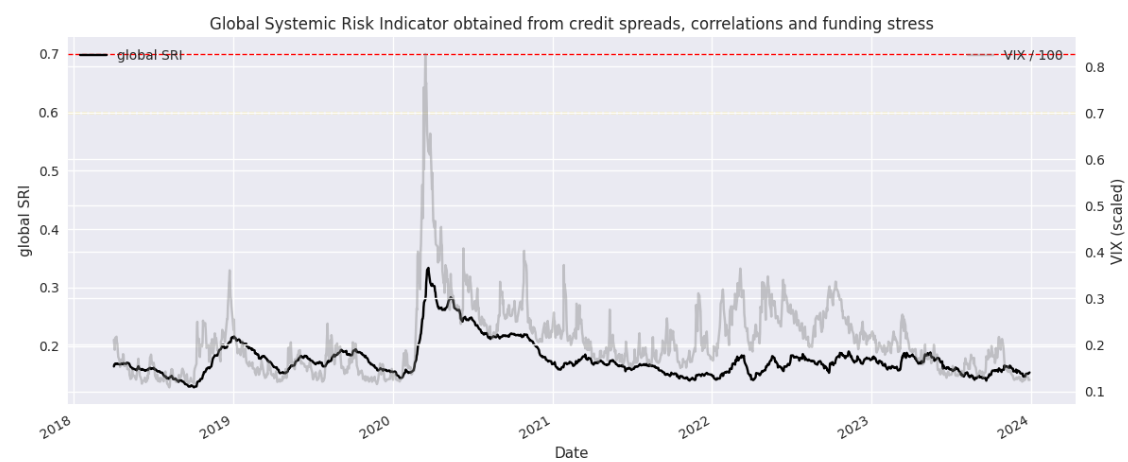
The result from the optimization indicate that the most important indicator is the one from student A based on the PCA monitoring framework

```
array([0.01869481, 0.        , 0.01203404])
```



Our integrated approach is able to signal abrupt movements in market volatility ahead of time, giving us the ability to prevent losses, eventhough it smooths some of the peaks a bit too much, so we miss small abrupt movements in the VIX when using our indicator. We can see our indicator breaks down at the end of the 2022 this is because the contribution of our DCC model based on the SOFR and Federal interest rate took a massive regime shift, one that it's historical as it changed the macro environment into one that had not been seen in decades (see the article by Irvine D). As we will see next allowing our model no lower bound on the importance of this SOFR and Federal interest rate indicator based on DCC gets us better tracking of the VIX but we loose our predictive power as the spikes are not seen ahead of time

```
array([1.        , 0.80739945, 0.03163214])
```



The change in the macro environment and the fact that a similar macro is not seen in the data gives our indicator some trouble, the good news is that the interest rates cannot be high for a long time, in fact, how long they have been high is what makes the current economical situation so unusual. According to Moore S. the possibility of the Federal reserved lowering interest rates next month sits at around 91% , those interest rates will perhaps return the correlation of our indicator into it's normal regime, thus allowing us to use our indicator as the correlations are likely to go back to their usual pattern once interest rates come down

0.3 Comparison with VAR model

In this section, we are going to compare our SRI with output prediction for ^VIX from a Vector AutoRegressive (VAR) model.

We are going to focus our attention to the period immediately before the spike in volatility that happened during the Covid crisis from approximately February to April 2020.

The crisis period February - April 2020 (included) will stand as our testing period. The training period will cover the two previous years, from April 2018 to January 2020.

The VAR model will be composed of

- the ^VIX index itself,
- the credit spreads as obtained from FRED (BBA10Y series),
- the TED spread from FRED,
- the SOFR - FFS spread, from FRED as well.

	^VIX	Credit Spread from FRED	TEDRATE	SOFR_FFS_Spread
2018-04-03	21.100000	0.0184	0.60	0.14
2018-04-04	20.059999	0.0185	0.64	0.05
2018-04-05	18.940001	0.0184	0.64	0.06
2018-04-06	21.490000	0.0184	0.64	0.06
2018-04-09	21.770000	0.0183	0.61	0.06
…	…	…	…	…
2020-01-27	18.230000	0.0204	0.25	-0.02
2020-01-28	16.280001	0.0204	0.23	-0.02
2020-01-29	16.389999	0.0205	0.25	-0.02
2020-01-30	15.490000	0.0208	0.22	-0.02
2020-01-31	18.840000	0.0213	0.23	0.01

[462 rows x 4 columns]

	^VIX	Credit Spread from FRED	TEDRATE	SOFR_FFS_Spread
2020-02-03	17.969999	0.0210	0.20	0.00
2020-02-04	16.049999	0.0209	0.20	0.01
2020-02-05	15.150000	0.0207	0.20	0.00
2020-02-06	14.960000	0.0205	0.19	0.00
2020-02-07	15.470000	0.0205	0.20	0.00
…	…	…	…	…
2020-04-27	33.290001	0.0324	0.72	-0.01
2020-04-28	33.570000	0.0325	0.65	-0.03

```
2020-04-29  31.230000                    0.0324     0.59              -0.03
2020-04-30  34.150002                    0.0323     0.47              -0.01
2020-05-01  37.189999                    0.0325     0.42              -0.02

[63 rows x 4 columns]
```

We fit the `VAR` model from the `statsmodels` package in Python.

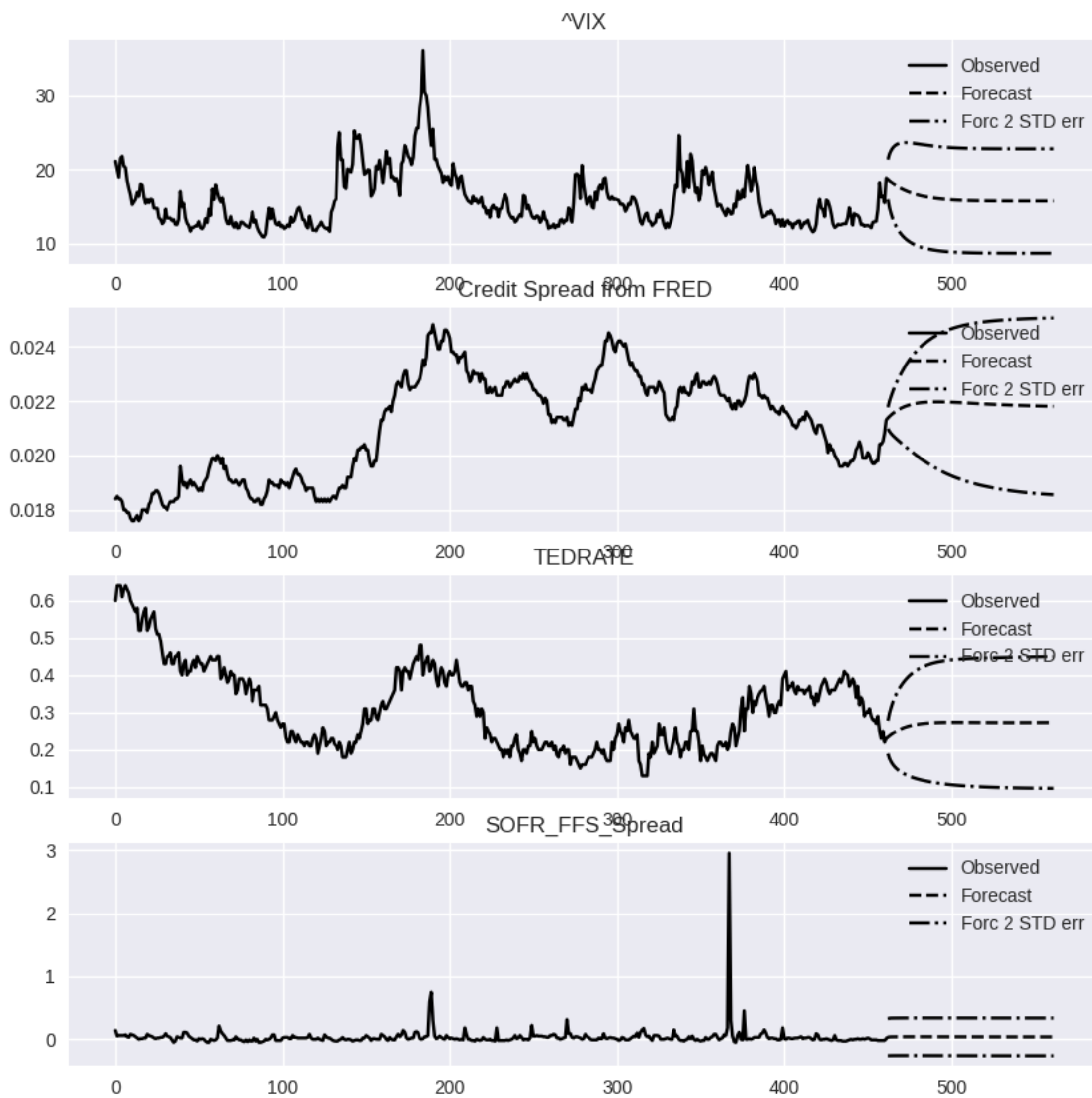
```
  Summary of Regression Results
==================================
Model:                         VAR
Method:                        OLS
Date:           Thu, 28, Aug, 2025
Time:                     22:59:46
--------------------------------------------------------------------
No. of Equations:         4.00000    BIC:                   -27.7384
Nobs:                     461.000    HQIC:                  -27.8471
Log likelihood:           3838.50    FPE:                7.50771e-13
AIC:                     -27.9177    Det(Omega_mle):     7.19064e-13
--------------------------------------------------------------------
Results for equation ^VIX
================================================================================
=============
                                coefficient       std. error           t-stat
prob
--------------------------------------------------------------------------------
-------------
const                              1.487511         0.880004            1.690
0.091
L1.^VIX                            0.927616         0.019953           46.490
0.000
L1.Credit Spread from FRED       -10.895717        40.989033           -0.266
0.790
L1.TEDRATE                        -0.390176         0.662717           -0.589
0.556
L1.SOFR_FFS_Spread                -0.127420         0.424356           -0.300
0.764
================================================================================
=============

Results for equation Credit Spread from FRED
================================================================================
=============
                                coefficient       std. error           t-stat
prob
--------------------------------------------------------------------------------
-------------
```

```
const                              0.000342         0.000129            2.646
0.008
L1.^VIX                            0.000015         0.000003            4.980
0.000
L1.Credit Spread from FRED         0.977149         0.006024          162.203
0.000
L1.TEDRATE                        -0.000262         0.000097           -2.689
0.007
L1.SOFR_FFS_Spread                -0.000100         0.000062           -1.601
0.109
================================================================================
=============

Results for equation TEDRATE
================================================================================
=============
                                coefficient       std. error           t-stat
prob
--------------------------------------------------------------------------------
-------------
const                              0.018327         0.014386            1.274
0.203
L1.^VIX                            0.000657         0.000326            2.014
0.044
L1.Credit Spread from FRED        -0.877149         0.670084           -1.309
0.191
L1.TEDRATE                         0.962961         0.010834           88.883
0.000
L1.SOFR_FFS_Spread                 0.016076         0.006937            2.317
0.020
================================================================================
=============

Results for equation SOFR_FFS_Spread
================================================================================
=============
                                coefficient       std. error           t-stat
prob
--------------------------------------------------------------------------------
-------------
const                             -0.131084         0.094941           -1.381
0.167
L1.^VIX                            0.001345         0.002153            0.625
0.532
L1.Credit Spread from FRED         6.118748         4.422205            1.384
0.166
L1.TEDRATE                         0.020679         0.071499            0.289
```

```
0.772
L1.SOFR_FFS_Spread                  0.206444         0.045783           4.509
0.000
================================================================================
=============

Correlation matrix of residuals
                               ^VIX  Credit Spread from FRED   TEDRATE
SOFR_FFS_Spread
^VIX                       1.000000                 0.382287 -0.016272
-0.033113
Credit Spread from FRED    0.382287                 1.000000  0.055383
0.032826
TEDRATE                   -0.016272                 0.055383  1.000000
-0.013576
SOFR_FFS_Spread           -0.033113                 0.032826 -0.013576
1.000000
```

Results are statistically acceptable. The model was directed to select autonomously the number of lag terms for the VAR independent variables.

```
VAR model was fitted with 1 lag terms
```

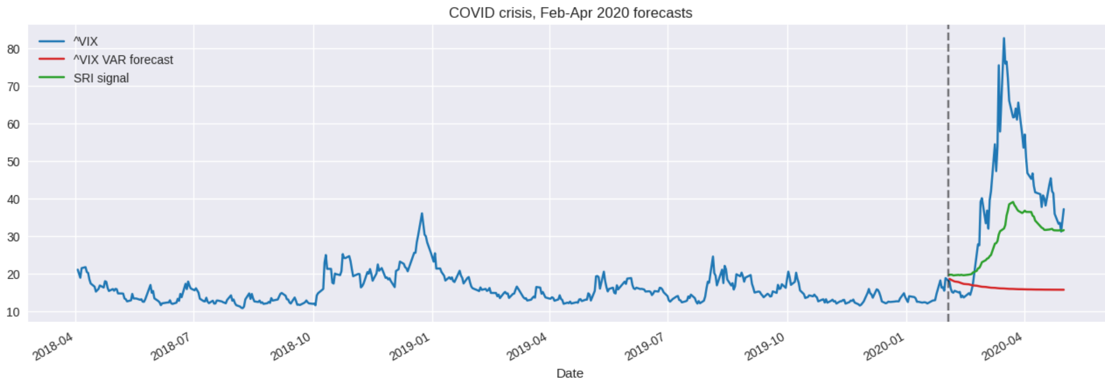


(461, 4)

2020-02-03	18.640695
2020-02-04	18.450905
2020-02-05	18.272151
2020-02-06	18.104234
2020-02-07	17.946625
	…
2020-04-27	15.751542
2020-04-28	15.749242
2020-04-29	15.747154
2020-04-30	15.745261
2020-05-01	15.743549

```
Freq: C, Name: ^VIX VAR forecast, Length: 64, dtype: float64
```

A graphic comparison of the output from the VAR model, the SRI, and the true path taken by ^VIX during the Feb - Apr 2020 testing period is presented below.



We see that the SRI performs better than the VAR model, although still quite not able of predicting the full intensity of the crisis.

0.3.1 References

- FRED database, **https://fred.stlouisfed.org/series**, accessed on 25 August 2025
- Molnar, C. (2025). Interpretable Machine Learning: A Guide for Making Black Box Models Explainable (3rd ed.). christophm.github.io/interpretable-ml-book/
- Afzal, F., Haiying, P., Afzal, F., Mahmood, A., & Ikram, A. (2021). Value-at-Risk Analysis for Measuring Stochastic Volatility of Stock Returns: Using GARCH-Based Dynamic Conditional Correlation Model. SAGE Open, 11(1). https://doi.org/10.1177/21582440211005758 (Original work published 2021)
- Arısoy, Yakup Eser, Quantile Regressions: Estimating Moments of the Stock Return Distribution (July 19, 2023). Available at SSRN: https://ssrn.com/abstract=4515136 or http://dx.doi.org/10.2139/ssrn.4515136
- Yahoo Finance. (2025). Historical data for ETFs was obtained from https://finance.yahoo.com/
- SciPy 1.0: Fundamental Algorithms for Scientific Computing in Python. Nature Methods, 17(3), 261-272. DOI: 10.1038/s41592-019-0686-2
- Irvine, D. (2024) A new era of higher interest rates, Forbes. Available at: https://www.forbes.com/sites/danirvine/2024/04/30/a-new-era-of-higher-interest-rates/
- Moore, S. (2025) Fed expected to cut interest rates, though inflation may be picking up, Forbes. Available at: https://www.forbes.com/sites/simonmoore/2025/08/16/fed-expected-to-cut-interest-rates-though-inflation-may-be-picking-up/