Introduction:

I had recently built a web facing app (link here) that uses a respository I made in GitHub (link here). The repo exclusively contains python code and works with the Streamlit API. There overall goal of the app is to run time series related analysis on financial time series.

The three functions the app can perform at the time of this publication.

- Historical regime – creating a historical Markov regime switching models and then create separate time series for each regime.
- Smoothed variance probability – calculate the smoothed probability of low, medium, and high variance of returns for a financial time series.
- Continuous wavelet transform – transforming those smoothed probabilities via continuous wavelet transform to create spectrum graphs

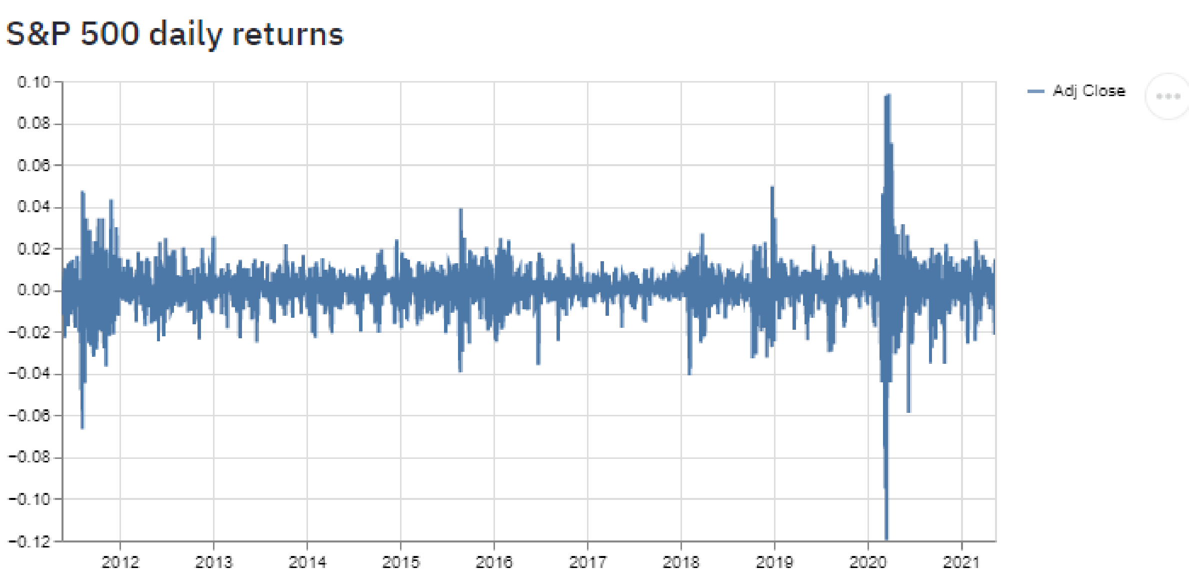
Methodology:

The data used in these experiments involves pulling data from yahoo finance by using the yfinance python API (see documentation here). The data was pulled at various time frames included (daily, weekly, monthly), and all of the prices are the Adjusted Close Price. The functions used for calculating the markov regime switching model is the python statsmodels API (see documentation here). The method used for finding the smoothed probability of the variance of returns for each regime is also built upon sample code from the statsmodels API (see 2nd example of the Kim, Nelson, and Startz (1998) Three-State Variance Switching link here). All of the code is accessible on the GitHub and all of the results are replicable using the Streamlit app.

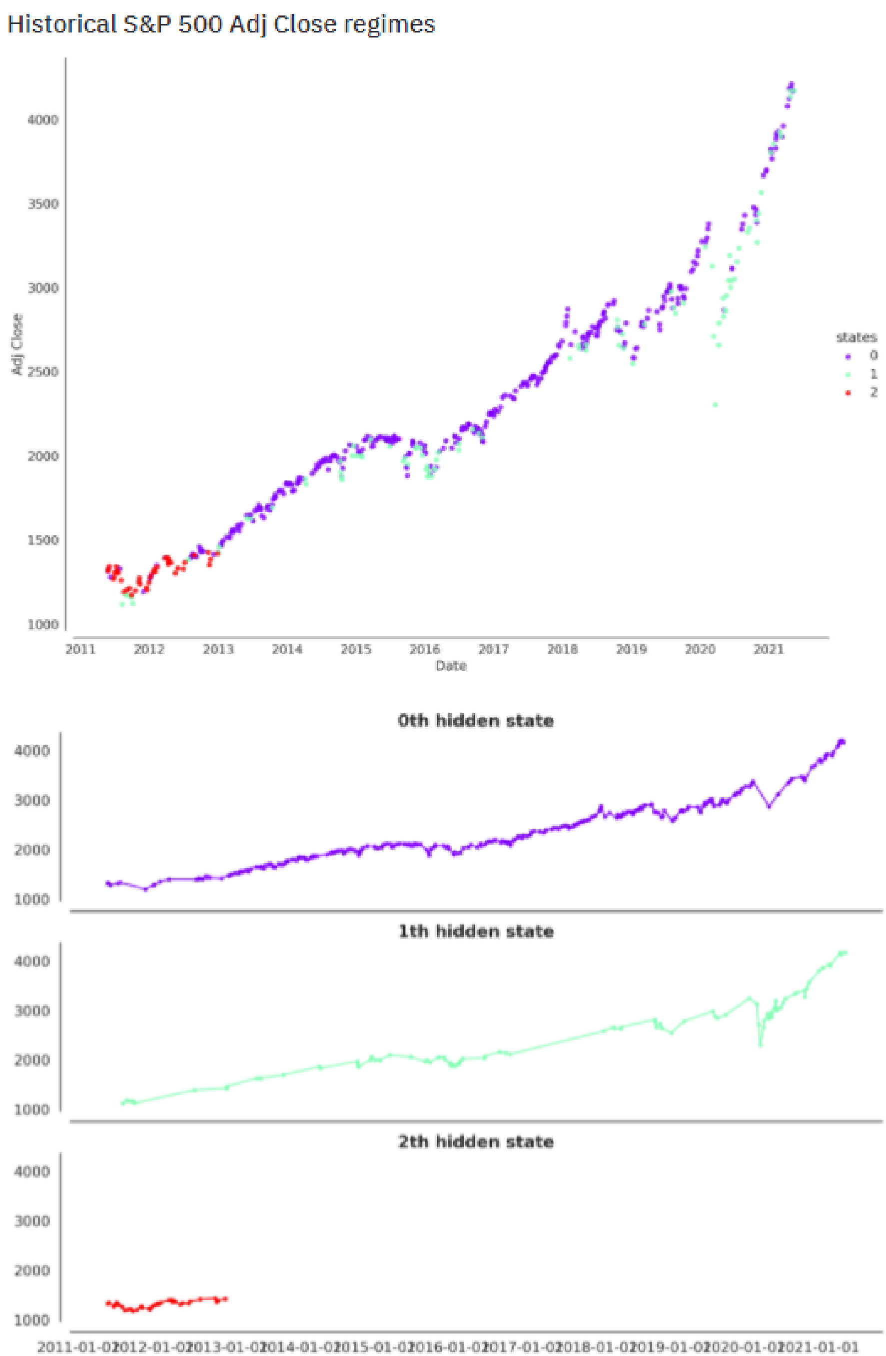
Background:

The smoothed variance probability is a Markov regime switching model. We first start with a security's adjusted close price, below is the adjusted close of the S&P 500.

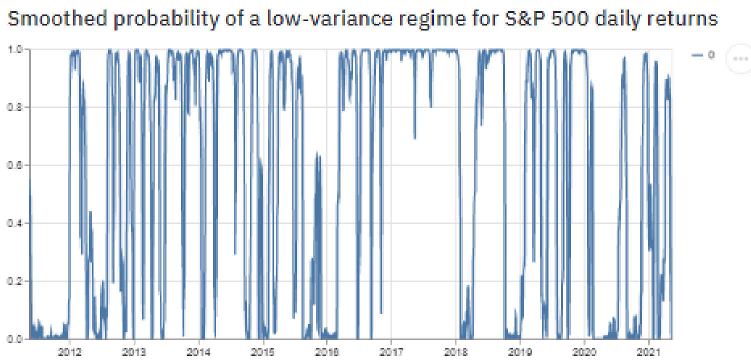
Then find the percentage change. In this case the prices are the daily returns therefore we will find the daily percentage change in price, which would look something like this.

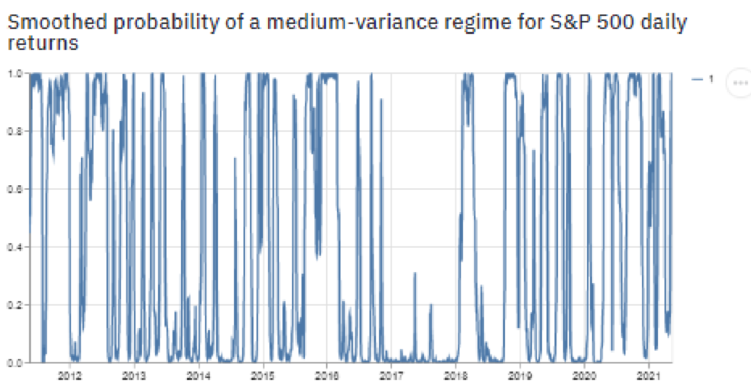


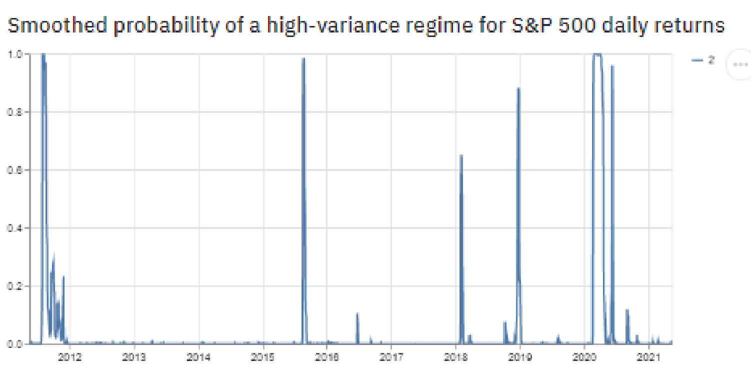
Then from there we construct a 3-regime Markov switching model dependent upon the variance of the returns. This is sort of parsing out the different timeframes of returns, and in this case, they are filtered by variance. Below is an example of a 3-regime model on price not returns, this is to give a visual idea (built in the app).



Then from there using the statsmodel method called smoothed_marginal_probabilities (see same documentation for example); that will provide three graphs that will show the probability of switching from one state to another. Finding the probability for switching regime for the daily returns of S&P 500 looks like.

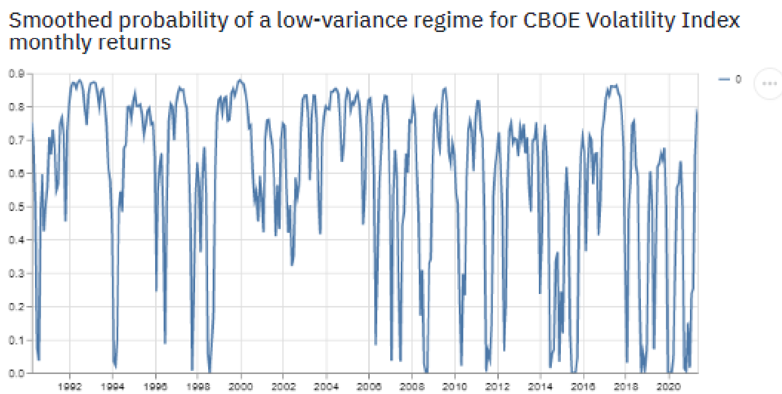


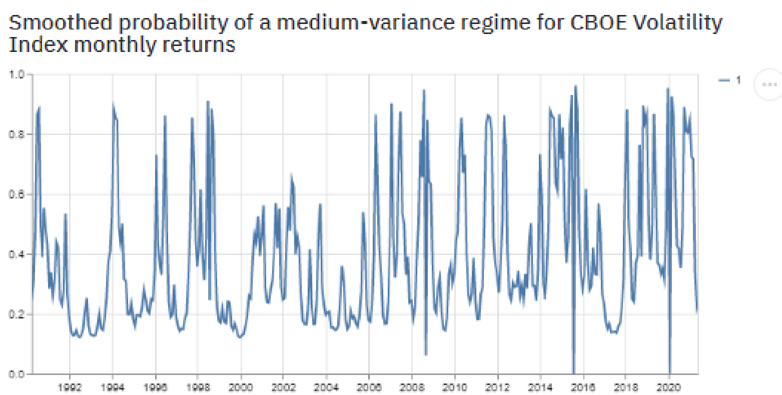


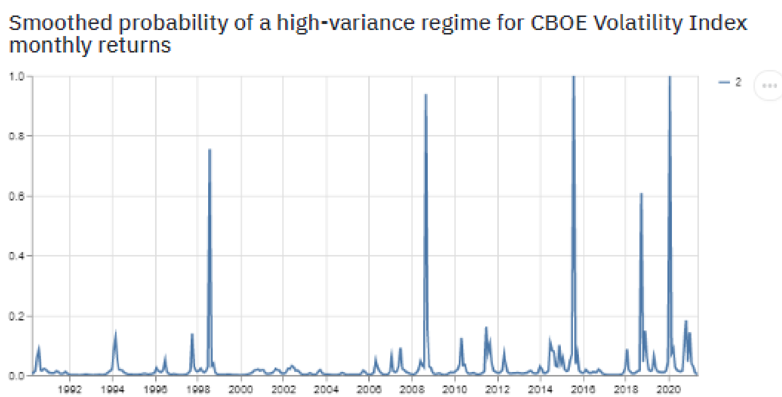


Results:

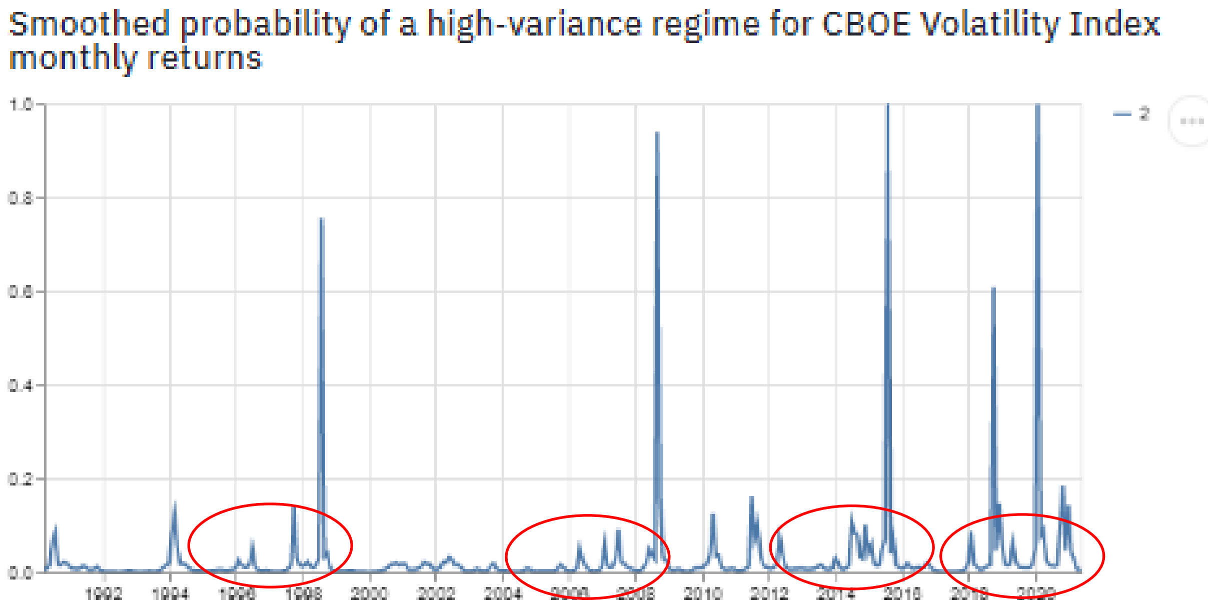
This test involves all of the historical data for CBOE VIX that exists. While searching through the smoothed marginal probability for the weekly CBOE VIX I noticed something in these graphs.





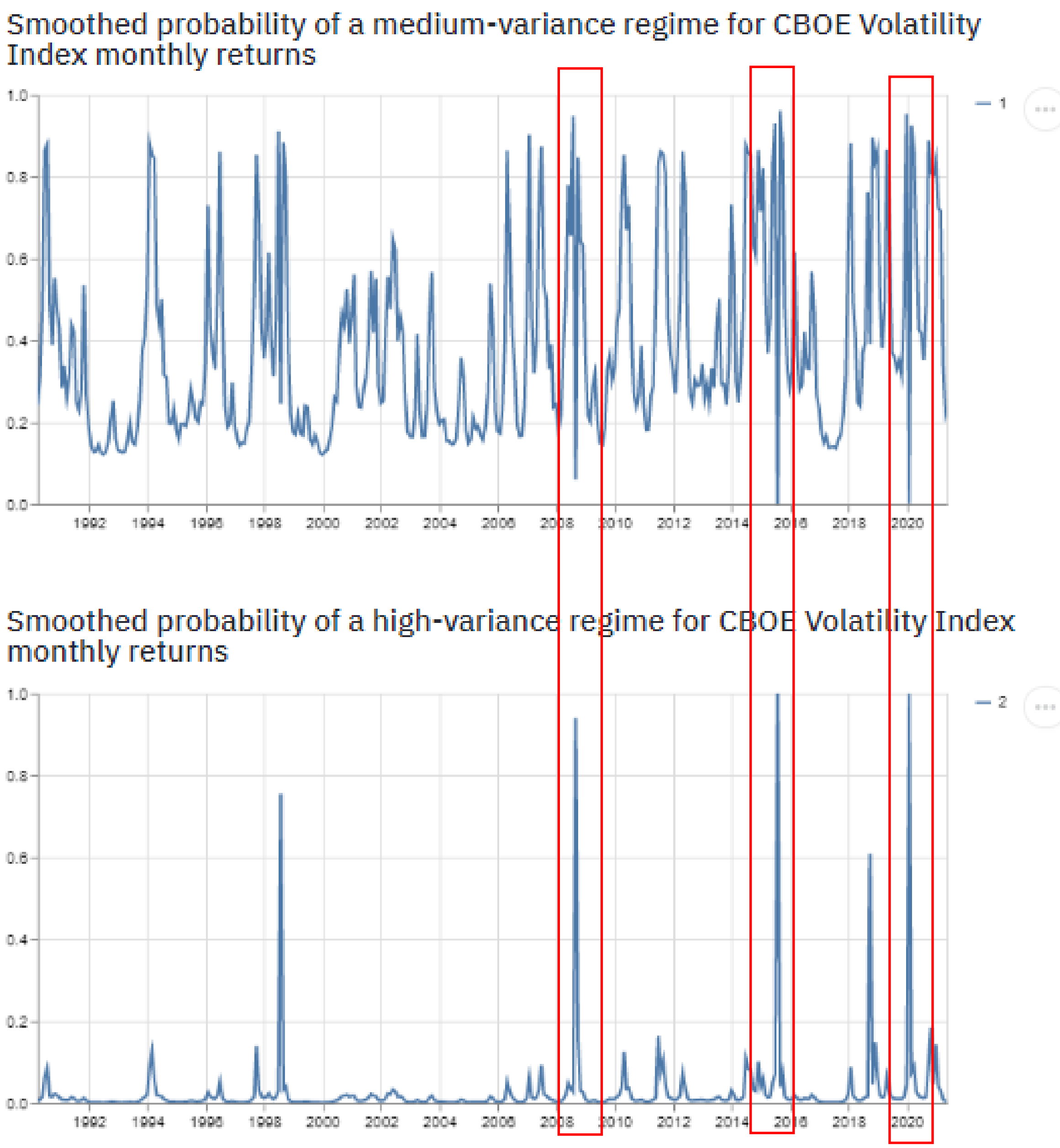


The first observation was looking at high-variance regime spikes.



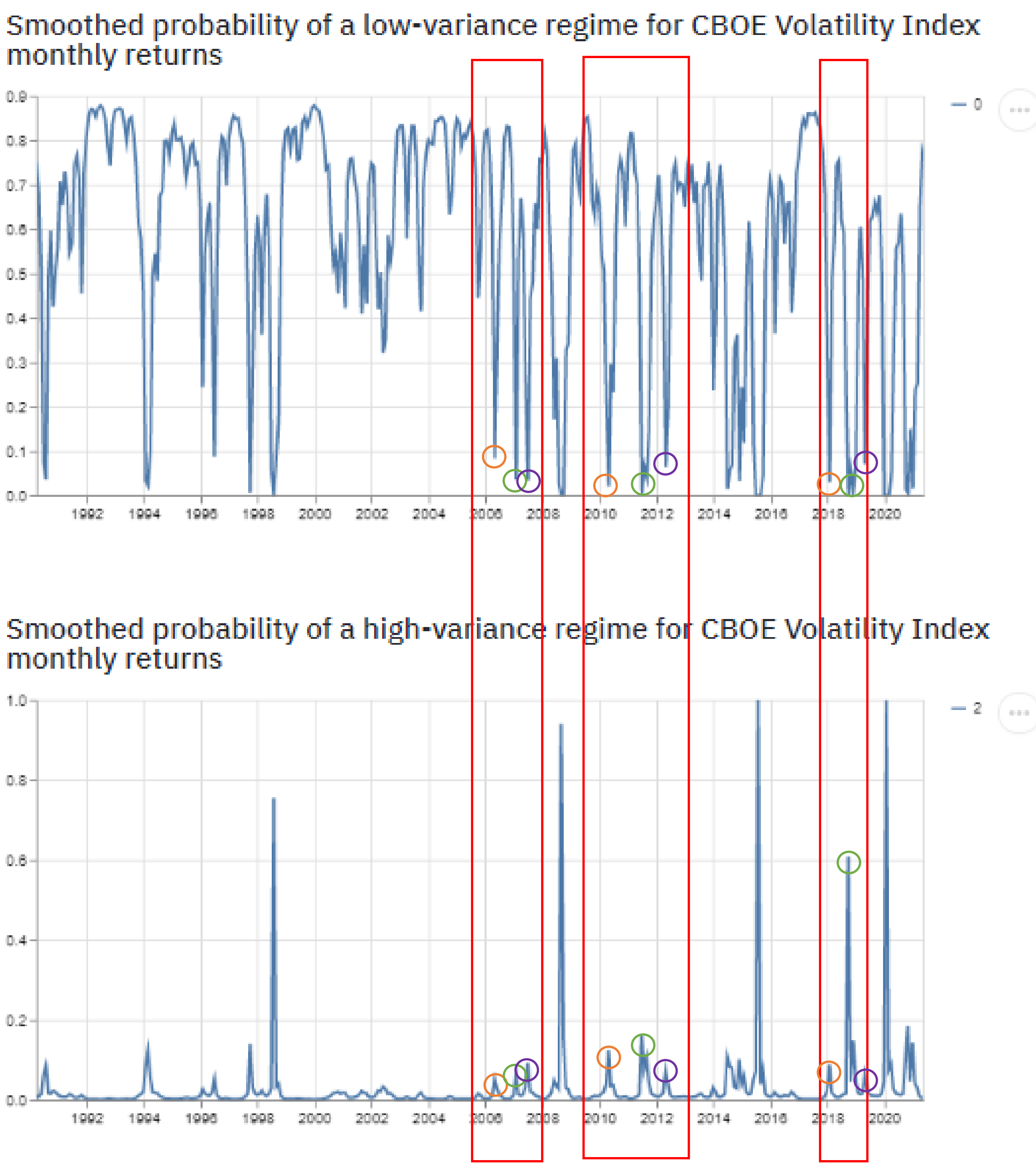
I noticed that there are some minor spikes prior to big spikes in high-variance regimes. This could possibly give lead to the idea that large spikes in high variance regimes don't occur by randomly they may be grouped with minor spikes.

The other thing that I had noticed was looking at the relationship between the two regimes. When looking at low-variance and high-variance we can see the change with large spikes.



We can see that the large spikes are associated with changes in the medium to high regime switches. This may give insight on the how these large spikes occur. It could be the case that the market switches from small to medium spikes which shows signs of more volatility and then they switch from medium to higher switches. It also shows that because most of the bigger spikes in high variance returns are associated with drops in the medium regimes there is already a *general consensus* of risk in the market.

When looking at low variance regime switches to high variance regimes switches we see.



This may mean that the large spikes only occur during transitions from medium to high variance, but more research is necessary. Something to investigate is the relationship with high-variance probabilities coming from either low-variance or medium-variance probabilities. Something to investigate is that a higher volatility may materialize to a bubble / recession if the transition comes from medium to high and it could be the case that small to high transitions will not materialize.

Ideas:

The initial goal of this was to apply this to left-tail hedging strategies. One possibility was looking at what happens before crashes. This still needs to be reviewed more by looking at the how the daily and weekly probabilities look for those specific dates. One idea would be to use the frequencies as lensing tools. It could be the case if we observe something on the monthly time series and then *zoom* in by looking at finer frequencies. Or it could be done vice versa going from daily to weekly to monthly.

Another idea is detecting whether higher variance regimes are market selloffs that will have rebounds, or they will materialize into bubbles. One approach would be detecting the medium to high variance changes, which looks like a sign of a bubble, and observing that scenario and then searching for the overvalued security or security class. The other method could be looking at value indicators (Buffet Indicator or Shiller PE CAPE indicator) and then linking medium to high variance changes to overvalue signals from those indicators. I could also make my own value indicators as well.

According to Black Swan Theory, Black Swans events contain a large amount of information within them. Looking at these smaller spikes (which technically are not Black Swans but sort of micro-Black Swans) they may contain important pieces of information within them. This could mean that looking into the spikes that are not associated to crashes to extract information from them.

This model could be used to predict left tails. Although this thinking opposes the theory behind modeling within Black Swan theory it seems like this model could provide an accurate predictor of Black Swans events. Even if the indicator had a *reliable* level of confidence in predicting Black Swans it probably could still be used, given that Universe (Taleb & Spitznagel) says that 95% of their left-tail hedges lose money (link here).

On the converse if you assume that financial models that dictate the market as well as the market itself does not have or promote antifragility then times of low variance would be times for finding cheaper left-tail hedges. There is some implications to this idea that lead to some problems. Problems that occur when using this model to find cheaper hedges include security type (inverse ETF, derivatives), duration on derivatives, and strike prices on those derivatives.

Areas of Concerns:

There is a large room for error. I break those areas into these categories.

- Programming errors – this could be errors within the statsmodels or within my code.
- Data errors – errors within the data that is getting pulled.
- Knowledge gaps – I do not fully understand markov chains, processes, or regime switching models to be fully confident in implementing a model. I also fully don't understand how to apply them in the python programming language let alone the statsmodels API.
- Statistical errors – this would be errors when drawing and making statistical inferences from the time series.

Another area of concern is the performing future analysis on this model. Markov regime switching models are a form of time series analysis. Now we are trying to analyze the time series of model that analyzes a time series. Possible next steps could be implementing more statsmodels tools or using another regime switching model, or continuous wavelet transform. But applying a second form of analysis may be overkill.

Another problem is that an indicator made from this may not be able to detect those high variance moments fast enough. But there are still applications of using machine learning to determine when a high variance probabilities will materialize to big selloffs.

Other uses:

- using it to for risk parity – volatility targeting models
- using it to stress test portfolios – determining what types of stress testing to use for a portfolio for a given timeline
- determining when rebalancing is necessary – if there are spikes that won't materialize then it probably isn't worth rebalancing and instead weathering
- building indicators – creating an indicator to detect changes in volatility or other financial time series
- applying this to different time series – market indices, interest rates, macroeconomic indicators, or portfolio

Diego Alvarez
Diego.alvarez@colorado.edu
LinkedIn | GitHub | Medium