VIX – An NYU CDS Capstone Project Proposal

Name: David Hamilton

Email: deh284@nyu.edu

I have over 15 years of institutional options trading experience on both the Buy and Sell Side in the US and Europe. Specific areas of focus in my career have included proprietary quantitative strategies (dispersion/correlation, relative value), market making and client facilitation.

This project will attempt to examine a highly visible, yet often poorly understood derivatives product using key elements of the CDS Masters curriculum including big data and machine/deep learning. Since it focuses on a relatively young, rapidly expanding area of the market, there is exciting potential for gaining valuable new insight(s) from research driven by data science.

What is the project?

The VIX index estimates expected volatility on the S&P 500 stock index. It does so by averaging the weighted prices of options on the S&P (puts and calls) over a wide range of strike prices. There are extremely liquid, exchange-based futures contracts based on the price of VIX as a means of trading and hedging said volatility.[1]

This project has both a primary and secondary goal. First, it will research and attempt to identify the possible existence of a short-term disconnect between the price of VIX futures contracts and the composite value of the underlying S&P options on which they are calculated. Second, if the initial study discovers meaningful inefficiency in the VIX/S&P relationship, it will look to design a realistic, actionable trading strategy that can profit from the potential arbitrage.

[1] VIX White Paper, Chicago Board Options Exchange (CBOE); 2015

What does the data look like?

A detailed description of the type of data that is required to address the problem. For example, will the students primarily be working with data aggregated from twitter feeds or other social networks, is it medical data, biological data, or financial data? Along with detailed information on the type of data there should also be context regarding where the data can be found. Will the organization provide the majority of the data? Is the data accessible via other avenues/ sources? Is an NDA required? How much of the data is available? Do the students need to gather data, model data, both, or neither?

The data required will come in the form of Trade and Quote (TAQ) datasets for all VIX futures and S&P options contracts to be used in the study. Ideally, these should reflect securities prices at the tick-data level. If that degree of granularity is not practicable, then the smallest possible increment of time (1-second updates, 5-second updates) that fits within the scope of the project should be used. The requisite datasets already exist and can ideally be procured from various resources within NYU. Otherwise, there are a number of external entities and financial data vendors who, given the nature of the project, could potentially provide the information for little or no charge.

What is the proposed scope of the project?

Is this a project for 1, 2 or 4 people? Is this a smaller well-delimited problem or the next step in addressing a grand challenge in your field? How many hours of work are anticipated for completion of the project?

Ideally this would be a project for 1 or 2 people. Its stated goal attempts to address two questions facing the industry: does the market currently price/trade VIX futures (and similar products) in the best way possible and, if not, are there new opportunities and potential for profit by taking a different approach? The proper investigation of both issues should require a minimum of 150-200 hours of work (75-100 hours/person in the case of 2 people).

What are the rubrics of success?

The Center for Data Science will be grading all of the final projects. In order to do this, we will need information from the stakeholders around what progress/ success would look like. For example, is progress classified as the creation of an algorithm that improves upon the current process? Each problem must be accompanied by a detailed description of the rubrics of success.

In other words, how will this project be judged and graded? What are the quantitative and qualitative metrics that can be used as metrics for successful completion of the capstone project. Since each problem would ideally propose a different set of solutions, these rubrics must be specific to each individual problem.

Successful completion for the initial phase of the project should be fairly easy to define. If the necessary data is compiled and researched in the proper manner, it should be readily apparent if there is a quantifiable disconnect that occurs in the VIX/S&P relationship and, if so, how often does it happen and to what degree (on average) when it does?

The second part, assuming the discovery of a tangible inefficiency, will be harder to evaluate. Since the only true test of success for a trading strategy is to demonstrate profits in live market conditions, any measures used to calculate the efficacy of project-generated ideas will be, at best, an estimate. The key will be in identifying how accurately real-life conditions have been simulated. In the instance of possible arbitrage profits, additional qualitative evaluation can be performed on the methodology used (probability & statistical analysis, machine learning, signal processing etc.) to devise the potential trading solution(s).