PHD RESEARCH AT ALGOLABS

December 12, 2018

INVESTIGATING THE USE OF ONLINE MACHINE LEARNING METHODS AS EFFECTIVE HEDGING MODELS.

1. Preamble

In this document, we will introduce the background and research questions surrounding the proposed PhD research theme of machine learning methods and their application towards developing effective hedging models.

2. Background context

The financial marketplace is a huge, ever-growing, complex and extremely exciting venue for trading all sorts of assets, such as Bonds, Equities and FX (Foreign Exchange). In this marketplace there are all kinds of players trading with each other, from individuals to hedge funds, investment and retail banks and even automated trading models. Transactions can be tiny in value (e.g. 10p) or huge (e.g. £100 mio)! To participate in the financial marketplace, you have to either buy or sell something. The main rule of making money in finance is to sell high, buy low!

AlgoLabs supports various brokerages who have clients that trade in the following marketplaces:

- Foreign exchange (FX, or Forex) and their derivatives
- Contracts For Difference (CFDs) based on:
    - All major world stock market indices (such as FTSE, DAX)
    - Energy (crude oil, gas)
    - Base and Precious metals (copper, gold, silver, platinum)

Central to the operation of a financial market are entities known as Market Makers (also referred to as Liquidity Providers). Market makers are businesses (such as an investment bank or a brokerage) that make / publish prices in a set of financial instruments (e.g. FX, indexes, commodities and so on). If a client wants to trade on the prices published by a market maker, the client must first place an order which is defined as a request to trade a given financial instrument. Market maker prices can either be ASK (sell) or BID (buy) prices. During the order process, clients can buy from market makers at the ask price and sell at the bid price. The difference between the bid and ask prices is called the spread. Market makers will compete for client liquidity by ensuring their spreads are as tight as possible, publishing ask prices which are competitively low and bid prices which are competitively high. If a market maker 'wins' a client trade, it then takes on the risk associated with that trade and assumes a position in the market (either 'long', 'short' or 'flat'). The market maker can choose to hedge its position, placing trades to flatten its position in order to (hopefully) make money. The market maker can do this by executing trades with tighter price spreads than those published to the clients or by using a specialised hedging model which makes more intelligent decisions about when and how to hedge.

3. The Research Problem

This PhD project focuses on the operation of an automated Market Maker. In this project we can think of an automated Market Maker as having three models fundamental to running the full lifecycle of the business (see also Figure 1):

1. The Pricing Model – the goal of this model is to win as much client flow as possible by offering competitive prices. For example, if the pricing model is able to intelligently respond to a rising market it will win more flow on the ask and vice versa.

2. The Hedging Model – the goal of this model is to make as much profit from client flow. For example a successful hedging model is able to decide an optimal time to place a hedging trade in order to maximise and lock in a profit

3. The Execution Model – the goal of this model is to execute market maker trades effectively and efficiently, with minimal transaction costs to external parties. For example, although Market G may have the better price, it may charge $10 per mio more than Market T

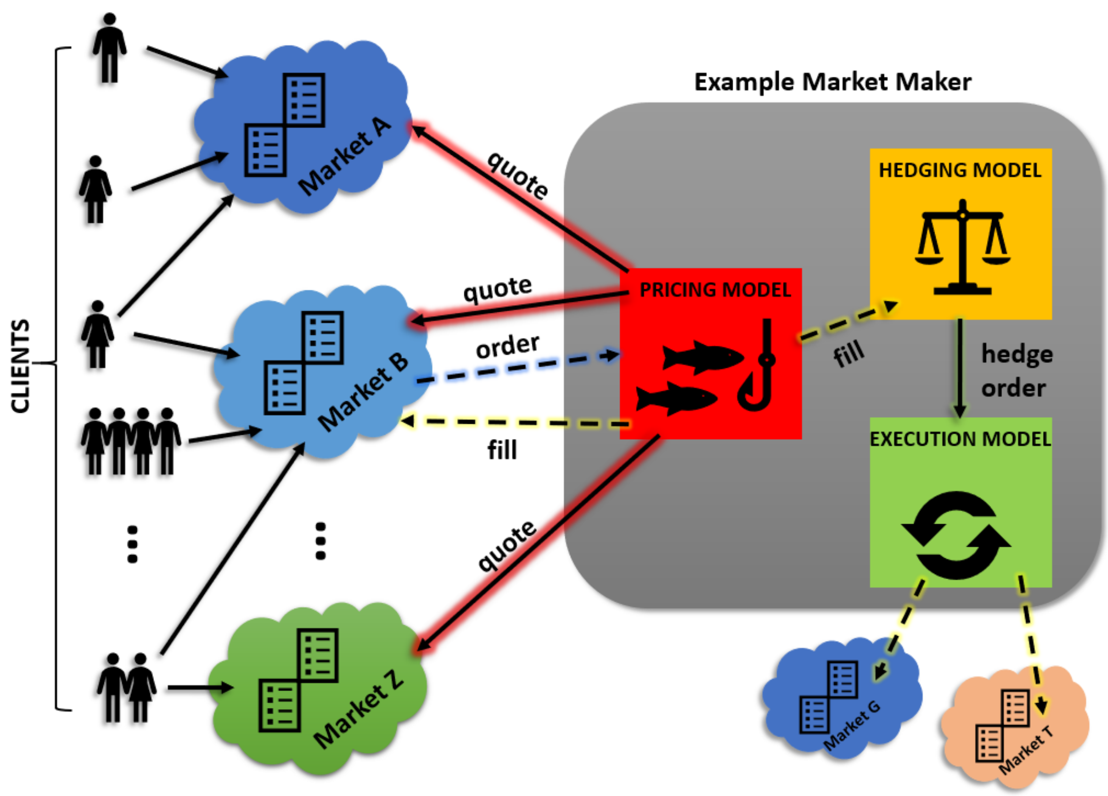

Figure 1. An example market maker, comprising Pricing, Hedging and Execution models. Liquidity in the form of client orders are continually being sent through to different market types. The Pricing model assesses ideal prices and offers competitive quotes with the aim of filling, or completing, the orders with. Upon winning client orders, the market maker assumes a position in the market and the Hedging Model can decide an ideal time to place a profitable hedging trade via the Execution Model and back in the market. Ideally the Execution Model works efficiently in order to minimise transaction costs.

Understanding the Hedging Model

A market maker's first priority is to construct a robust "quote" (bid and ask price pair) to its clients, which it can re-quote many times a second. The use of online machine learning techniques to construct prices of the market maker's quotes is to be investigated in a similarly-related PhD project that this project will work closely alongside with.

The main focus of **this** PhD is to look downstream at how we can effectively manage risk making decisions in the Hedging model. The Pricing Model sends quotes to various markets, which then get passed onto end clients which can be either:

- Retail clients: individual small clients manually clicking on a GUI
- B2B (Broker to Broker) clients: aggregated business from one broker to another, which in turn could be from retail or institutional.
- Institutional clients: these are more professional clients with larger deposits, they typically have accounts with multiple brokers and will trade via their own platform or programmatically via a FIX API connection.
- The market maker will 'win' flow from a variety of these clients, depending on how aggressive its pricing is. Clients send over orders to the market maker, which can then either:
- Fill completely (client asks to buy 1 mio EUR/USD, market maker sells 1 mio EUR/USD to them)
- Fill partially (client asks to buy 1 mio EUR/USD, market maker sells them 500k of it)
- Reject (do not fill anything, market maker must give a reason why)

The Hedging Model only needs to be concerned about trades that are filled or partially filled as it is these trades that represent the risk taken on by the market maker. Then the main and ongoing consideration of the Hedging Model is in accumulating the market maker's "position" from the resulting fills. The concept of position in the market is illustrated in Table 1 below and Figure 2 overleaf.

Time	Price	Side	Amount	Symbol	Client Name
2017/11/06 13:44:03.211	1.1123	Sell	1,500,000	EUR/USD	A
2017/11/06 13:44:03.211	1.1126	Sell	500,000	EUR/USD	C
2017/11/06 13:44:03.211	1.1129	Buy	1,000,000	EUR/USD	A
2017/11/06 13:44:03.211	1.1132	Buy	6,000,000	EUR/USD	Y
2017/11/06 13:44:03.211	1.1135	Sell	2,000,000	EUR/USD	Z
2017/11/06 13:44:03.211	1.1133	Buy	4,500,000	EUR/USD	Z

Table 1: A sequence of different EUR/USD trades managed by a given market maker for four clients (termed A, C, Y and Z). These trades are plotted in Figure 1 to illustrate the market maker's differing position over (a very short space of) time.

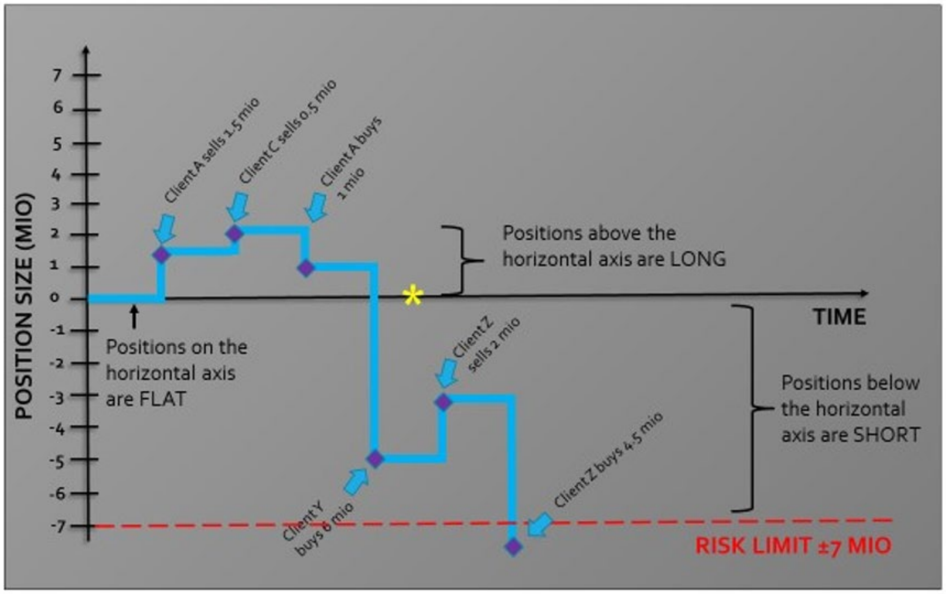

Figure 2: The vertical axis shows the changing position size of a given market maker over time and according to the sequence of client trades detailed in Table 1. The market maker's position is categorised as being FLAT, LONG or SHORT if the position size is on, above or below the horizontal axis, respectively. Position changes are equal and opposite to those of the client fills. All market makers have a risk limit which in this example is +/- 7 mio. If the market maker breaches this limit, so-called 'hedge' trades will be placed to move its position to zero (this is sometimes called 'flattening').

Defining Risk

In our context risk is defined as:

1. The position you build up in the market (summation of client trades pushing the market maker longer or shorter) and,
2. The PnL (profit and loss) of the positions you have built up (dependent on the prices that the client trades have been executed at – further decided 'upstream' by the Pricing Model)

Resources are finite so keeping the two elements of risk in check is important. This can be done by placing "hedge" trades that mitigate the risk by bringing net position closer to zero. One can imagine a disastrous scenario whereby the market maker has a very large position in the market but not enough PnL to reduce its position to mitigate further loss. Therefore it is essential that position risk limits are put in place. Aside from avoiding position risk limit breaches, knowing when, what and how much to hedge is the job of the market maker's Hedging Model. It is worth reminding the reader that we can only hedge if we have a long or short position in the market, we cannot hedge if we are flat.

Unrealised and Realised PnL

PnL, or Profit and Loss, can be of two types – realised and unrealised:

- Unrealised PnL is also called 'floating PnL' and it refers to the market value of a financial instrument (e.g. EUR/USD) at a given point in time relative to its original price. This value may go up or down.
- A realised or 'real' PnL occurs when you take steps to flatten your position by buying or selling, to lock in (and realise) a market price and associated PnL for that instrument.
- You will profit if you buy at a price lower than what you originally sold at, or if you sell at a price higher than what you originally bought at.
- Conversely, you will take a loss if you buy at a price higher than what you originally sold at, or if you sell at a price lower than what you originally bought at.

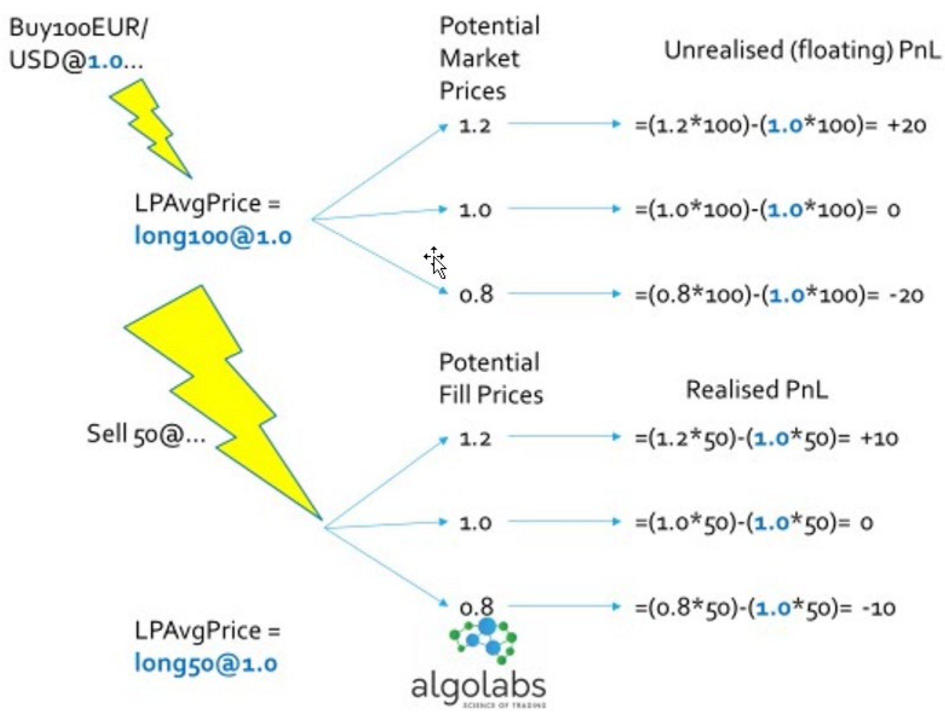

*Figure 3: The example illustrated above shows how unrealised, or floating PnL is calculated, and how realised PnL is calculated. Let's say we are a liquidity provider (LP) and after buying 100 EURUSD at 1.0 (represented by the smaller lightning bolt) our position is now **long100@1.0**. The market is offering a range of prices for us to sell EURUSD at, from 0.8 to 1.2, and we can calculate our potential, floating PnL from these by comparing the price at which we bought to the prices which we could potentially sell at. Now, we decide to take steps to flatten our long position by selling (so becoming shorter – see larger lightning bolt). We could sell at different prices again, ideally, we want to sell at a price higher than what we bought at. If we sell 50 EURUSD at 1.2 we will make a profit of 10. Conversely, if we sell 50 EURUSD at 0.8 we will make a loss of 10. Note that even after selling, our position remains long (long50@1.0).*

The charts below in Figure 4 are based on real data and show how prices, market positions and realised and unrealised PnL values are all related over time. The lines representing the average bid and ask prices from LPs are very dynamic, reflecting frequent changes in price as orders are placed across the market. The average client price line focuses on the prices traded by the clients and is less dynamic as clients do not place orders on every single price that is published by an LP! The position chart shows whether a client is long or short for a given order price, and how much by. Positions above the zero line are 'long' indicating that the client has bought more than sold, whereas positions below the zero line are 'short' indicating that the client has sold more than bought. Looking at the average bid, ask and client price lines in conjunction with the client position line should indicate how successful a particular client is. For example, as prices increase around timepoint 08:09:56, the client's position is short 2 mio, whereas it ideally would be long (remember: buy as price increases, sell as price decreases). This is also reflected in the client's realised PnL decreasing around the same timepoint.

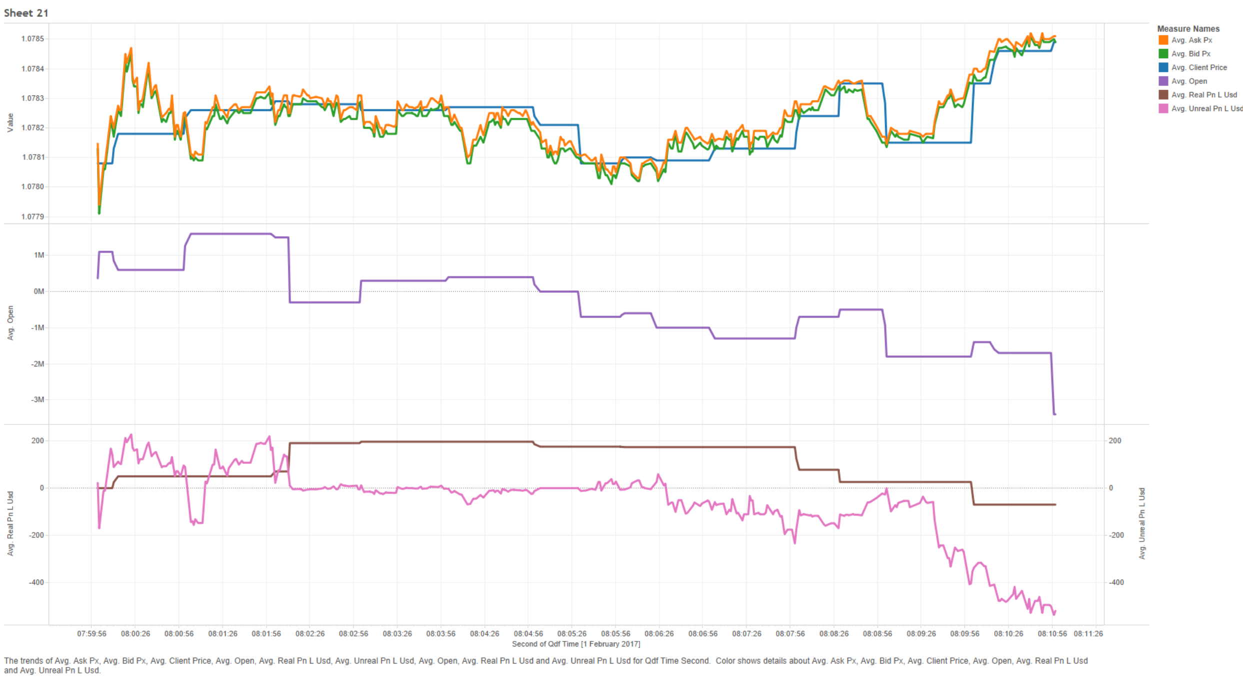

Figure 4: How prices, positions and unrealised and realised PnL values are related over time

The unrealised PnL chart is just as dynamic as the average ask and bid price charts because it represents PnL values that could be potentially realised with each change in price received by the market. Unrealised PnL also reflects the client's success in trading in the same way that the realised PnL chart does.

A- and B- Booking

There are effectively 2 default hedging strategies currently employed by all brokerages – 'A booking' and 'B booking':

A booking – off the back of every client trade that the market maker executes, the market maker goes back to the Liquidity Provider (LP – who quoted the prices used in the trade) and places a trade of an equal amount and opposite side with this LP. For example:

1) Market maker publishes a EUR/USD Ask price of 1.15 – this is an inflated price offered by LPs (which is 1.14)
2) Client X places a Buy trade with market maker, buying 1000 EUR/USD at rate of 1.15. At same time, Market maker goes back to LP and places a Sell 1000 EUR/USD trade at rate of 1.15.
3) The market maker instantly makes a profit (1.15-1.14)/1.15*1000=$8.69

See also Figure 5 which illustrates the concept of A booking.

The rationale for A-booking focuses on speed since we don't want the LP to requote at a less favourable price. A-booking offers a "riskless"[1] pass through with the broker taking a small cut in the middle with the inflated price. The profit made on this is almost guaranteed (depending on how fast the LP fills and whether they reject). Brokers typically do this for what they deem "smart" clients where they consistently make money or build up too big a position for them to handle. The market maker assumes no position or unrealised PnL; all PnL is all realised in milliseconds in small positive amounts based on the mark-up added to the client's prices.

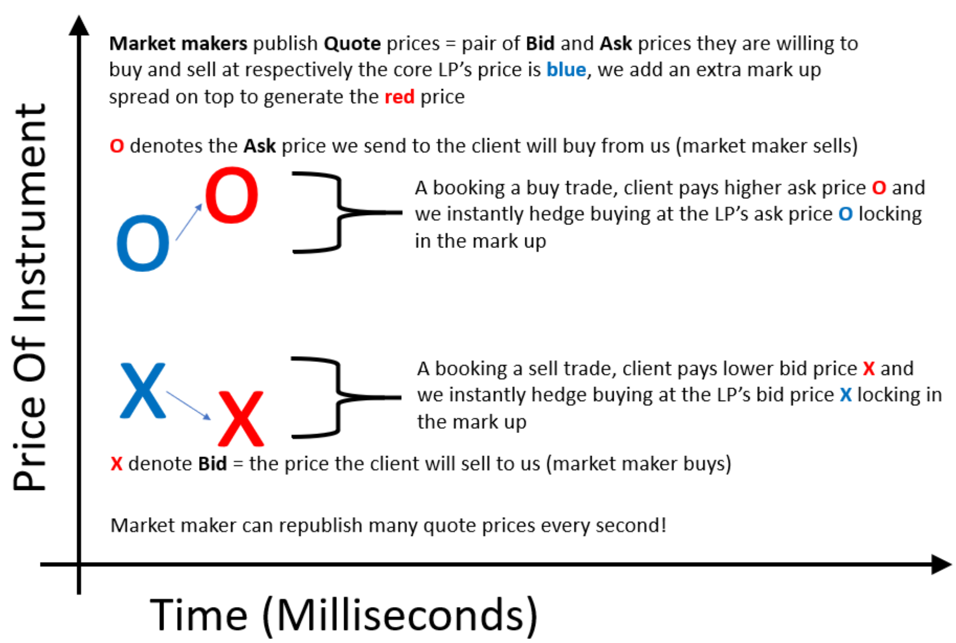

Figure 5: The concept of A-booking

[1] *Provided that the client does not default on this risk e.g. by going bankrupt. In this scenario, the broker would take on the risk entirely.*

B booking – this is the opposite end of the spectrum to A booking, and this hedging strategy is typically used for "stupid" clients who consistently lose money, so in theory no hedges need to be placed. With this model you have unrealised PnL at the mercy of the clients, you only realise PnL when they reduce or flip your net open position.

In reality both hedging approaches are overly simplistic, no client is consistently 'good' or 'bad' at trading, so a Hedging Model may employ a hybrid approach, for example:

- the hedging model will place hedge trades when it predicts that the market will turn against, mitigating any unrealised PnL swings
- if the position is within the market maker's global risk limits and predicts that its position is favourable i.e. where the market is trending, the Hedging Model will not place any hedges or lift existing hedges

4. Proposed Aim and Problems of PhD Project

The aim of this PhD project is to investigate the use online machine learning methods as effective hedging models. Data from historical client trades and their corresponding prices will be used to conduct these investigations. The performance of any online hedging strategy can be compared against a default Hedging Model currently in place which is essentially is a rules-based model with global parameters that specify maximum net open position as well as take profit and stop loss boundaries to keep these realistic. A simple market direction indicator based on price momentum is also used to decide if the positions are favourable or not. The default model also uses different risk overlay parameters by running back test simulations on historical data, using Genetic Algorithms to search the parameter space effectively and allowing visualisation of the risk trade-off when running certain parameter combinations.

The key challenges to the research project that we foresee are as follows:

- The Hedging Model must be relatively fast, capable of managing the flow of over 200,000 trades each day and 300 mio price updates from LP's.
- Although we know there are 3 main types of client flow: retail, B2B and institutional, they all have different patterns of trading, different biases on symbols, different times of day when they trade the most etc. Can we cluster these differences and if so is it better to mix the flow and hedge in aggregate or create smaller models hedging the flows in isolation, which is more effective/ profitable?
- Although it is possible to simulate performance (backtest) by replaying the historic tick data we have recorded (over 10 years across 70 symbols), there is a natural feedback loop which cannot easily be quantified (often referred to as Market Impact Analysis) -, if we hedge excessive amounts over a short time period we will affect the prices, how can we factor this into the model design?
- To maximise profitability, short term market direction forecasts need to be made to decide whether our position is favourable or not, but what time horizon should be used to make the forecasts over – seconds? minutes? hours?
- We know that price forecasting is a hard problem (some even argue impossible), there are indeed huge levels of noise in the underlying data, can we factor this uncertainty in the forecasts into how we look at our position and decide whether to ride the risk or to hedge?

- How can we factor in Economic News Releases as a main driver of price volatility in the FX market? The release date and time (and even a predicted figure and market impact factor too) of these macro-economic indicators are known in advance, how should we hedge around these releases? The current market default is to reduce risk or even flatten arbitrarily 5 minutes before the release, but is this optimal? Can we intelligently ride risk based on the expectation of the number to be announced?
- In the FX market natural triangles are formed between currency pairs, e.g. EUR/USD - GBP/USD - EUR/GBP etc. Should we care about the risk represented by each currency pair within their respective triangles? Or is it better to aggregate currency pair risk into a single currency risk e.g. EUR, USD and GBP currency positions. If we deem the position too risky which of the different pairs should we use to hedge?
- Relating to the other market making PhD project that you'll be closely working alongside with, is there another way to exit out of risk instead of going to the outside market (executing via another LP, thereby paying their spread)? Maybe we could somehow skew our prices to clients to naturally flatten positions? How and when should this be done and how-to co-ordinate effectively with the main pricing model?