[XTX]
MARKETS

HOW LONG BEFORE YOU KNOW IF AN ALGO IS ANY GOOD?

MAY 2021

The role of (good and bad) luck is commonly underappreciated in the world of algos. Many traders will try an algo five times and form a pretty strong opinion based on what they see.

Is this reasonable? Let's find out.

THE EXPERIMENT

To make things simple we focused on a single currency pair - GBPUSD. We chose three order sizes - 10mio, 50mio, 100mio.

We randomly selected 10,000 order start-times over the last few months. Then we flipped a virtual coin to decide if each order would buy or sell.

Finally we assumed we got 20% of all aggressive fills on the CME from the start-time of each order until the order was completely filled. A simple VWAP.

Bear in mind several things:

- These executions never happened.
- There was therefore zero change in supply/demand dynamics.
- There was therefore zero market impact.
- There was therefore zero information leakage.
- This algo pays spread.

Overall, we'd expect to see some small slippage to arrival price on average to reflect the spread we paid. However, the executions are going to look pretty great overall since the orders to be hedged were imaginary and therefore didn't impact the market prices.

By this we mean the CME traded naturally - in the absence of these imaginary orders, which would've tilted supply/demand dynamics - and for the experiment we just assumed the algo could magically have received the fills it wanted.

To be clear this particular algorithm is imaginary and does not exist. It is purely a thought experiment, designed to show that even an algorithm that performs objectively well over the long run will have extremely noisy outcomes over a small sample of orders.

RESULTS

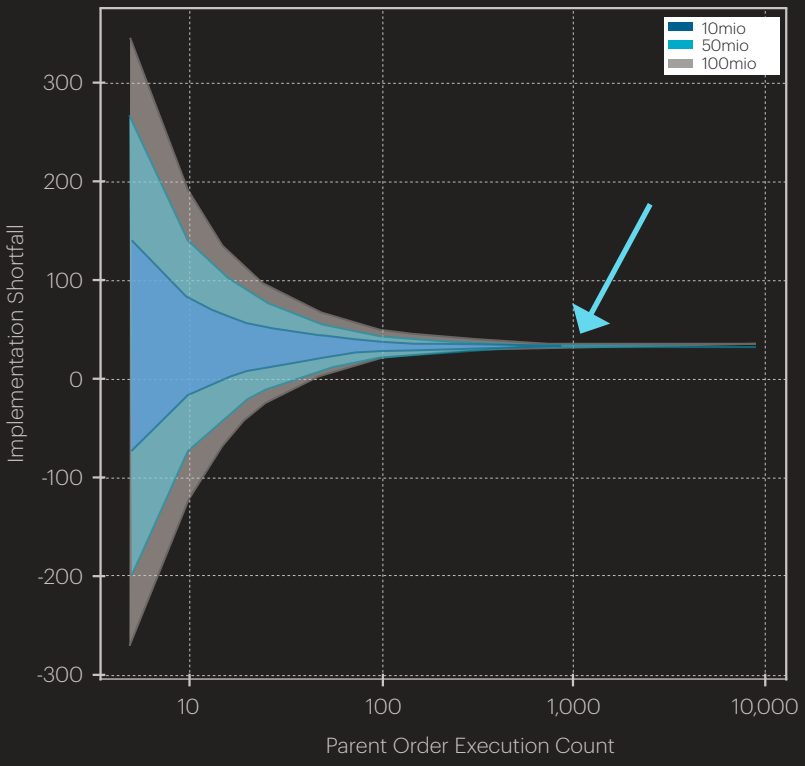


Chart produced by XTX using CME market data

The first thing to review is highlighted with the big blue arrow. This chart tells us that in the long-run the average implementation shortfall of this algo converges to about $40/mio, which is about half a pip in GBPUSD. That is what we'd expected to see:

"Overall, we'd expect to see some small slippage to arrival price on average to reflect the spread we paid. However, the executions are going to look pretty great overall since the orders to be hedged were imaginary and therefore didn't impact the market prices."

Performance figures start out pretty noisy, especially on the large 100mio orders which take longer to fill, allowing the market to randomly drift more during each order.

However, by the time we reach our blue arrow at 1,000 orders, we can clearly see the average implementation shortfall across all runs converges around $40/mio.

Nice! $40/mio is equivalent to about half a pip in GBPUSD. Clearly that is a superb hedging cost for 50mio or 100mio GBPUSD.

RESULTS AFTER FIVE RUNS

How does it look after five runs? Super noisy. The error bars (standard error of the sample mean) for the 50mio orders are highlighted with blue arrows below.

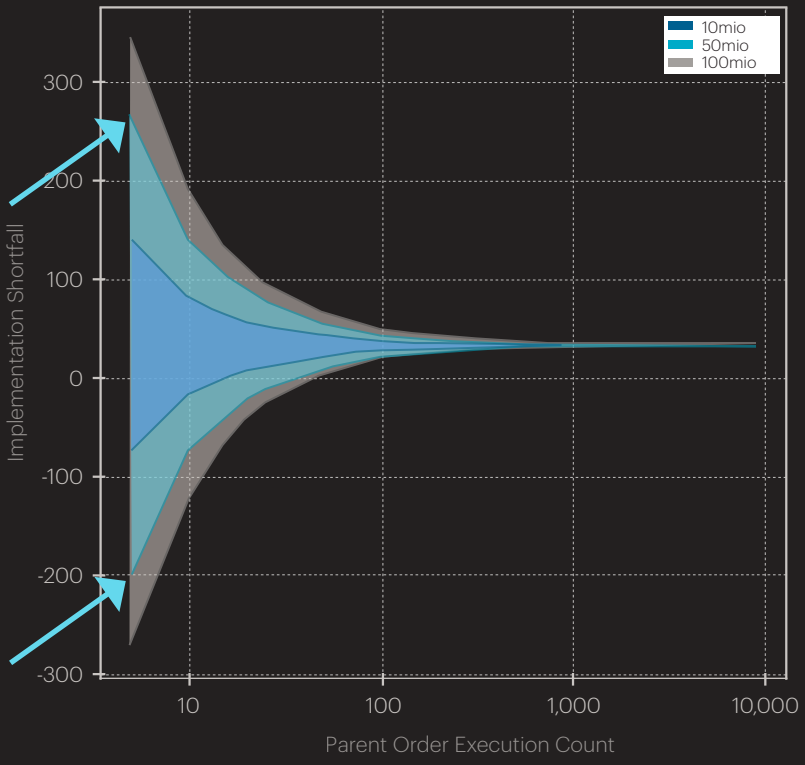


Chart produced by XTX using CME market data

After a single ticket of 50 million you may easily think the algo beats mid by 2.7pips (-$200/mio) or has slippage to mid of 3.8pips (+$280/mio). The difference in experience is extreme.

You can see the error bars for different combinations of order sizes and number of runs in the chart. The error bars for each size are colour-coded; simply move along the x axis to see how much they shrink for a given number of orders. Individual runs are

extremely noisy but even after five orders of 50mio we do not know much more. After five runs we can see the error bars shrink a little … but only to minus $180 (beat mid by 2.5 pips) to plus $220 (slippage of 3 pips).

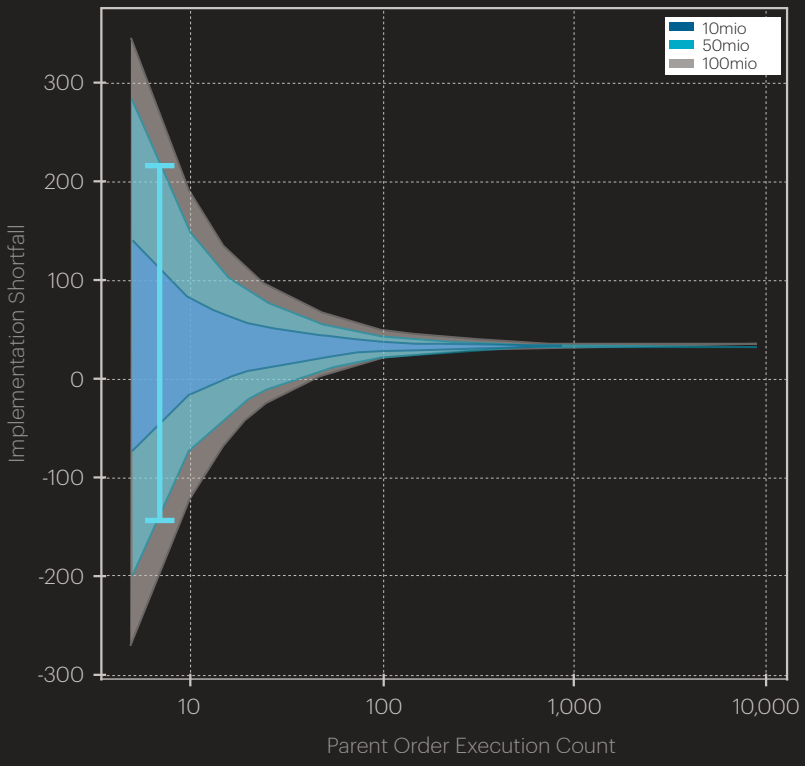
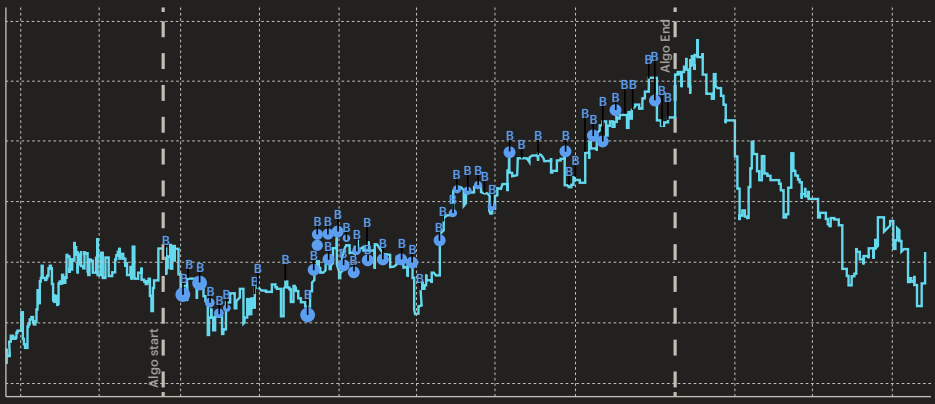


Chart produced by XTX using CME market data

The first five orders may therefore easily look great, terrible or average and still tell us almost nothing about this algo. As pattern-seeking animals it is really hard not to interpret a tiny sample set of five orders and try to deduce a pattern.

This is especially true because pure chance will often give us quite intuitive-looking price action charts on individual runs where we can spot trends, reversions, market impact and so on. Charts like the below are fairly common.



Illustrative image (not real data) showing a buy order whose start and end time appear to coincide with a drift higher in market prices

It is so hard not to intuitively rationalise these charts. If you saw this you'd surely agree it looks like the algorithm is causing impact.

However, let's not forget this hypothetical algorithm has objectively strong performance over the long run: it only costs half a pip on average to trade 50mio GBPUSD. Even so we will sometimes encounter evocative chart patterns. The crucial point is to remember this will happen and not leap to a conclusion based on a chart of a single run.

SO HOW MANY ORDER RUNS DO YOU NEED?

As we reach 100 executions a clear picture emerges for the smaller 10mio orders, where the trader now knows the average expected outcome to +/- $10/mio.

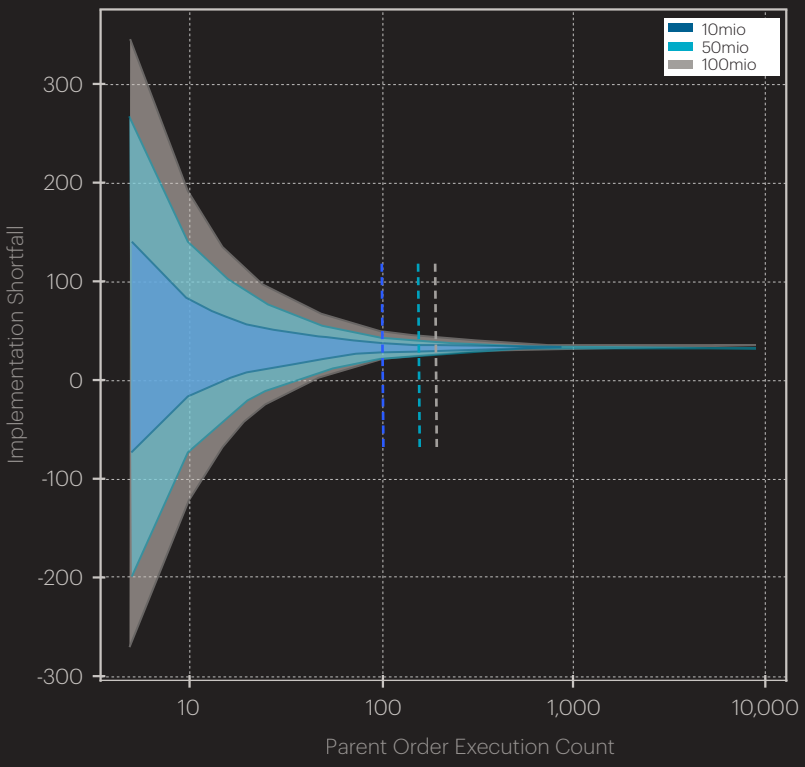


Chart produced by XTX using CME market data

We mark that with the blue dotted vertical line. For 50mio orders you require about 200 runs and for 100mio orders around 300 runs to reach the same level.

The reason we need more runs for larger orders is that they take longer to complete so the market can drift more in this time. This adds more noise. It is the same with less liquid and more volatile pairs.

The more orders you do, the smaller those error bars become and the better you can see the characteristics of an algorithm.

HOW TO GET ENOUGH DATA

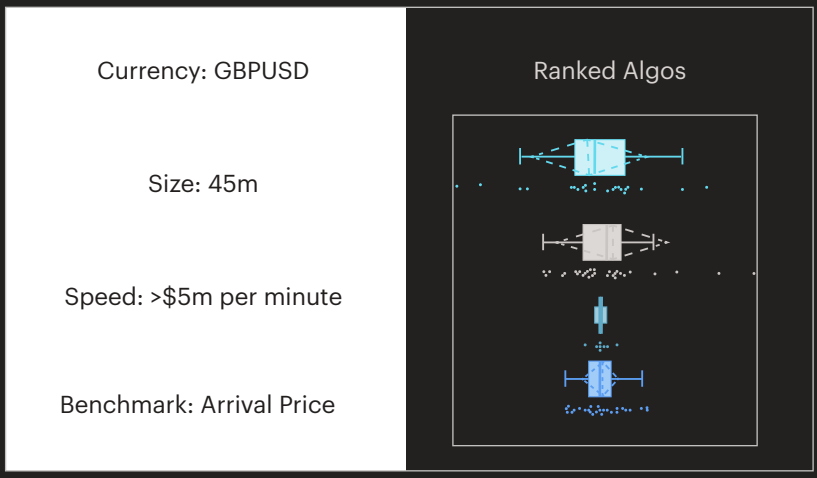
The rule of thumb seems simple enough: you need at least 100 runs of each algo in each pair. However since in real life we must normalise results by controlling for things like time of day, conditions, speed of execution, and parent order size ... this means you probably need more like 500+ runs of each algorithm in each pair.

That is completely impractical for any single client. No-one has enough orders.
The solution is independent Transaction Cost Analysis. These providers can create peer universes.

Almost all of the popular independent FX analytics firms offer this product today.

The idea is that many clients opt-in to a shared universe of metadata. No client can see anyone else's orders or sensitive details. This remains private. However, all opted in clients can see aggregated results.

For example, they might look at the implementation shortfall of all algorithms in GBPUSD of around 50m for the month of March across a sample of 675 runs.



Illustrative image (not real data) of a peer universe comparison of algorithms

When all the results are aggregated, the noise is reduced and the good algorithms float to the top of the results and poor algorithms to the bottom.

Even better - a client may see if a particular algorithm performs well without having to try it for themselves and pay away performance whilst learning. If an algorithm improves that will be visible, too.

TAKEAWAYS

• Know that you'll be tempted to form far stronger opinions than the facts can support for a small number of observations. Be alert to this and actively guard against this natural psychological bias.

• If you don't already then sign up to use peer universe tools to filter out candidate algos that are worth trying. Previous results on a large sample of orders are as good a guide as exists. Work with independent Transaction Cost Analysis providers to improve these tools and make them more useful for the buy-side.

• Nonetheless overlay some sense of intuition. The problem with peer universe data is that other people's circumstances don't exactly match your own. You may have faster investment alpha than average so need to trade faster. Or you may heavily customise an algo such that it produces different results for you than others. The data will point you in the right general direction but still requires a dose of good judgment on top.

• Whenever you feel tempted to judge an algo after five runs - it happens to us all - please think of this study. Recall that this hypothetical algo with objectively strong results over the long run - it buys 50mio GBPUSD for 0.5 pips - is likely to deliver an average result after five runs of between minus 2.5 pips and plus 3 pips. There is simply too much noise or 'luck' in which outcome you'll get over a handful of runs.

• The single biggest performance advantage you can get as a trader is more data to help you select the right tool for the job.

IMPORTANT NOTICE

This Document is issued by XTX Markets Limited ("XTX"), which is authorised and regulated by the Financial Conduct Authority of the United Kingdom (the "FCA"), with FCA FRN: 711945. XTX is a private limited company incorporated in England & Wales, with company number 09415174. XTX's registered office and principal place of business is R7, 14-18 Handyside Street, London, N1C 4DN, United Kingdom.

This Document is issued by XTX only to and/or is directed only at persons who are eligible counterparties for the purposes of the FCA Rules. This Document must not be relied or acted upon by any other persons (including, without limitation, persons who are retail clients or professional clients for the purposes of the FCA Rules). The distribution of this Document may be further restricted by law. No action has been or will be taken by XTX to permit the possession or distribution of the Document in any jurisdiction where action for that purpose may be required. Accordingly, the Document may not be used in any jurisdiction except under circumstances that will result in compliance with any applicable laws and regulations. Persons to whom the Document is communicated should inform themselves about and observe any such restrictions.

This Document is for information purposes only. The Document does not constitute an offer to transact in, or the solicitation of an offer to transact in, securities, derivatives, FX transactions, or other financial instruments, in any jurisdiction.

Although the information in this Document is believed to be materially correct, no representation or warranty is given as to the accuracy of any of the information provided. Certain information included in this Document is based on information obtained from sources considered to be reliable. However, any projections or analysis provided to assist the recipient of this Document in evaluating the matters described herein may be based on subjective assessments and assumptions and may use one among alternative methodologies that produce different results. Accordingly, any projections or analysis should not be viewed as factual and should not be relied upon as an accurate prediction of future results. Furthermore, to the extent permitted by law, neither XTX nor its employees, directors, officers, shareholders or service providers assumes any liability or responsibility nor owes any duty of care for any consequences of any person acting or refraining to act in reliance on the information contained in this Document or for any decision based on it. Past performance cannot be relied on as a guide to future performance.

Any reference in this Document to a counterparty trading with XTX (or XTX trading with counterparties) is a reference to XTX providing liquidity quotes through certain specific platforms, and all trades will be settled or cleared with or through XTX's prime broker(s) or CCPs, and not with XTX directly.


[XTX]
MARKETS