# Algos and Egos

# 2017 The Year Financial Markets Ate Themselves

Robert Hillman, 22 December 2016

This is part of a series of notes on a rapidly developing theme in financial markets and investing – namely the collision of algos and egos. In this note I write a fictional review of how a future observer may look back on 2017.

Nothing written below is true.

So 2017 is over. But according to the Chair of the Federal Reserve Nigel Farage, "we ain't seen nothing yet chaps". With global electronic trading exchanges suspended for the last six weeks while we await the launch of the new franchised London Metal Exchange open outcry pits, we take a look back on 2017 and try and piece together the events that led us here at what seems like breath-taking speed. Despite no shortage of political and economic shocks throughout the year, financial market volatility was almost entirely driven by the interaction of algos and egos.

The first sign of trouble was the seizure of the copper futures market leading up to the expiry of the front contract in November. There had been some warning signs. Volumes had been increasing in the last few years and the proportion of positions held by traders designated as speculators had reached a high of nearly 60% of total open interest at the end of 2016. While this trend had not gone unnoticed it had been interpreted as a sign of health, associated with a deeper and more liquid market providing the hedgers and consumers of physical copper with more efficient pricing and lower transaction costs. We now know this not to be the case.

## Volume explosion

What we know now of course is that the trading activity in copper futures, as we have subsequently observed across almost all electronically traded instruments worldwide, was carried out by a small number of entities who aggregated and automatically executed orders from millions of 'customers' worldwide.

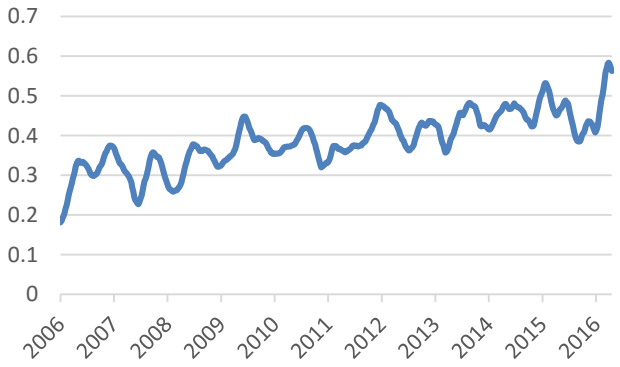


Figure 1. Total Managed Money Positions as a % of Open Interest, Copper Futures (Source CFTC & Neuron)

The chart shows a 12 week moving average of the weekly proportion of total managed money to open interest.

We use the term customer advisably because according to the just released preliminary findings from the recently initiated SEC-CFTC Commission, only 0.002% of the orders appear to have originated from an account that can be linked directly to a human. The rest were traced to two hedge funds in New York and one in Oxford, UK. To put in some perspective, on November 14th between 08.00.000 and 08.07.023 EST there were some 9 billion orders received by the exchange before systems shut down.

## A new type of hedge fund

The hedge-funds in question Quantizapp, Digidoodle and Spuggle were all created in 2016. The term hedge-fund itself may not be entirely appropriate. The business model of these entities was based on that of the crowd-sourced hedge funds that first appeared in 2015. This new breed of company ran competitions in which entrants were challenged to forecast streams of data. Entrants knew nothing of the exact nature of the data except in so far as they were encoded streams of market data. Competitors could use whatever techniques they wanted to analyse and forecast these data, and had merely to supply their algorithms in the form of replicable code to the competition host. The host firm then selected the winning entrants and deployed the algorithms in real-time with investor capital. Trading initially focused on futures and single-stock equities, but by 2017 had spread to ETFs, emerging market bonds and equities, credit-default swaps and European power derivatives. A percent of the profits accruing from trading was awarded back to the competitors in the form of bitcoin, or competitors could invest their winnings in the fund. At first these competitions were run on a quarterly cycle, but in early 2017 the second wave of competition-funds started a more continuous version of the game, with daily competitions facilitated by the introduction of straight-through-processing from submission to deployment, entirely removing humans from the process. It is thought that the whole operation of Spuggle is supported by 3 people, one of whom is a full-time yoga instructor and nutritionist.

## Malthusian data dynamics

Exactly what happened in the weeks following the launch of these funds is unclear but here is our attempt to understand it. It appears that certain competitors, in an attempt to increase their chance of winning, submitted multiple entries. But they did not stop at two or three. By automating the process of data-downloading, algorithm creation and submission, by August some were allegedly submitting several thousand algorithms a day. It is understood that initially some restraint was applied in order to not spook the competition host or overwhelm their systems. But this quickly unravelled as each entrant attempted to gain a temporary advantage on the others. While these companies had at first been dealing with a few hundred crowd-sourced algorithms they suddenly found themselves with several million at a time. It soon became apparent that the statistical techniques on which the host's business model was founded (using colourfully named statistical learning techniques such as 'bagging' and 'boosting') were ill equipped to deal with such a high number of entrants, and such a small set of examples over which to evaluate them. The hope of discovering useful trading algorithms disappeared virtually overnight. As one unnamed ex-employee from Quantizapp put it - 'we went from a talent seeking exercise which felt at least as plausible as the X-Factor to one with about as much success as The Apprentice".

## Front running

As multiple rounds of the games were played and overlapping supporting data sets were supplied, it became increasingly easy for entrants to decrypt the data and to figure out which underlying markets were being traded. Several of the encryption techniques used were open source and embarrassingly easy to crack. It is believed that there may have been some attempts by professional hedge funds to submit systems that they subsequently front-ran by trading ahead of them during the real-time trading phase. But it wasn't illegal or necessarily unethical behaviour that generated the real difficulties which were to follow.

## Fake news

As the competition hosts began to perceive that their data encryption techniques were being picked off they attempted to counter it by supplying fake data to competition entrants. This was deemed perfectly reasonable as a means of improving the chance of discovering genuine patterns in the real data. Some argued at the time it was merely an application of stochastic resonance. It is believed that during August 2017 Digidoodle began supplying exclusively fake data, hoping to disrupt the dynamics they believed were

occurring. In a statement on their website the firm claims that "Trading in a post-truth world requires post-truth data. While our data is not the cheapest, it is guaranteed to be completely uncorrelated to any known market data that has ever existed. Or your money back".

The introduction of encrypted fake market data into the trading ecosystem marked the beginning of the process that led us to what people today are calling the Great Data Inflation.

## Professional Investors Awake

Some hedge funds already specialised in big-data began feeding competition data into their systems, and discovering – unsurprisingly perhaps – that there appeared to be some predictive value combining the by now easily decrypted competition data sets and published performance data. As entrant identity was tagged and both their submission, evaluation sample, and real-time out-of-sample performance data were openly available to track, this led to the launch of second generation funds, in which competitors attempted to predict the winners of each first generation competition. It is thought that first generation competition hosts were some of the biggest clients in these second-generation funds as they attempted to hedge exposures to their own algos.

Just before the November market seizure, investment bank Bernstein Brown announced the launch of an ETF (ticker FFS Equity) based on a proprietary index they had been publishing for select clients over the previous few months. The derivative promised to give retail investors access to a liquid derivative linked to the strongest performing winner of each competition, marketed with the strap line "Learn how to trade like the best, but without the homework!".

## Code Mining

And it wasn't just market data that traders were mining. Sometime around late Summer the distinction between data and code vanished. Entrants built their algos using open source code components (predominantly but not exclusively written in Python and Julia), and many of them supplied their code back and forth into GitHub and other popular code repositories. Now GitHub itself became a source of data. Traders downloaded code from GitHub and data from the competition hosts and recombined them to create hybrid algorithms. These were placed back into GitHub. And the blurring worked the other way too. The encryption techniques used algorithms to create the streams of market data, requiring only a single keyword to initiate the algo which created the data stream. Code became data and data became code.

Even the perception of code has changed, a fact quickly recognised by some of the world's smarter investment consultants. As recently graduated Harold Smithers, analyst at Padlock Partners gatekeepers to 6 trillion dollars of US pension plan assets puts it "we slam the door in the face of any investment manager who comes to us bragging about how many million lines of code they've written. We don't want coders who can code anymore. We don't even want managers. We just want uncorrelated quant – pure and simple".

## Enter the gamers

Around late Summer a new breed of competition entrant was established. In the past few years, researchers at AI company NeuralSim had noticed that in training their machines to win in certain games of skill such as Go, it was helpful to create simulated games in which the machine could play itself. In effect learning by doing, bootstrapping its own knowledge, rewarding successful behaviours and punishing mistakes. What was found extra useful in these exercises, and particularly in situations where there was there was excess noise, was to filter the learning environment. To alter or augment reality. One of the oldest and most recognised examples of these techniques is to use blinkers on a race horse, or to use a head-up display to 'optimise' the visual sphere of a fighter pilot in order to downplay distracting features and enhance those more important.

## Translated Reality ®

Combining augmented reality, learning-by-doing and behavioural research on framing by Kahneman and Tversky, NeuralSim created the concept of Translated Reality. The idea - like most of those in the recent revival of artificial intelligence - was incredibly simple. Sometimes people are better at performing tasks when they aren't trying or thinking too hard about what they are doing.


Figure 2. Trading Dollar-Yen (Source Wikimedia Commons[1])

The image shows a trader during the training process. He is trading USDJPY during May 2015, has had a profitable run and is beginning to reduce his long US Dollar exposure (as he applies his brakes) in anticipation of a Bank of Japan meeting. A few seconds later he suffered catastrophic losses when the yen suddenly appreciated on the back of comments by Bank of Japan governor Kuroda expressing concerns with the rate of yen depreciation. In subsequent rounds the trader quickly learnt to reduce exposure quicker by feathering the brake pedal (taking risk off in modest clips to manage slippage).

NeuralSim took this idea further by creating video games in which players had to drive a car as quickly as possible through what appeared to be a computer generated randomly changing terrain. However, the terrain and circumstances in which gamers drive was actually a 3D representation of a real financial market. What was over the horizon or around the next bend was governed by what financials markets themselves were doing, see Figure 2.

Trending prices were represented as straights, choppy mean-reverting markets as randomly spaced chicanes, and rapidly accelerating price bubbles as blind humps in the road. Fluctuating degrees of liquidity and market resilience were represented by varying the road conditions, and anticipated data-releases and known option-maturities were represented as congestion on the road ahead. Drivers' responses: direction, speed, acceleration and so on were translated into trading strategy parameters such as size and sign of position. The drivers general level of aggression was measured and mapped directly into the urgency parameter of an implementation shortfall execution algorithm. What would have been a possibly overwhelming deluge of streaming financial market data to a trader, was filtered, distorted and translated into an entirely different experience.

Players could solve one problem while thinking they were tackling another. And the next step was obvious and followed almost immediately. Once sufficient skill was recorded (in practice no more than a dozen human players recorded a few hours of driving), NeuralSim was able to apply its learning algorithms to the recorded simulation data to super-charge the learning process. Recognising that the company already had thousands of hours of actual driving experience as a result of its research into self-driving cars, the logical conclusion, and where we are today, is that machine learning algorithms originally designed to control self-driving cars have become the predominant participant in commodity futures markets.

[1] By Derzsi Elekes Andor (Own work) [CC BY-SA 4.0 (http://creativecommons.org/licenses/by-sa/4.0)], via Wikimedia Commons

## Droning on

According to some conspiracy theory websites, this practice is not new. There are some who believe that the US military has contrived with the gaming industry for several years to embed reality into video games. There is an allegation that on occasion real (but digitized) feeds of Iraq conflict zones have been fed into the live online playing of Call of Duty. This approach has some support on ethical grounds, its proponents arguing it is a logical and more powerful extension of the blank-cartridge in a firing squad. Actual drone pilots in Texas are thought to be in fact providing training data for Amazon drone delivery services.

## Uncertainty Principle

Regulators have their work cut out. Having only recently decided that source code needed to be filed with the SEC, within weeks the distinction between code and data disappeared. And possessing the source code itself was arguably always useless. The complex and nonlinear nature of some of todays' execution algorithms means that even if regulators had an instantaneous copy of source code and the market data on which the code operates, it would still be impossible to retrospectively unambiguously recreate the output. Source code itself is seen as an outdated concept by many of today's execution specialists. They argue this issue is nothing new and point to the fact randomization has been used to disguise trading footprints created by execution algorithms for several years.

## Learning to spoof and spoofing to learn

A further problem for regulators is deciding what is permissible behaviour. One of the behaviours that the translated-reality gaming algos picked up on and exploited was bluffing. It turns out that this method of behaving so as to fool competitors in order to gain an advantage, or to induce the revelation of information, is intrinsic to almost all human (and animal) strategic interaction. But in financial markets the similar practice known as spoofing (the submission of orders with no intent to trade on them) is illegal.

It is not hard to see how a new framework for regulation may be required rather quickly. When algos are being created and deployed faster than a human can physically oversee them, arguably one of the key reasons for using machine learning in the first place, the only solution appears to be oversight by an even faster machine.

## Confirmation Bias

And so here we are today. It is unclear exactly how events unfolded but around late October we started seeing some very unusual behaviour in markets. The most visual indication was exponential growth in open interest and volumes of transactions. This caused a knock-on effect on the more traditional systematic hedge fund industry, dominated by CTAs, trend follower funds and risk-parity index products. As volumes grew in commodity futures markets many of these funds automatically moved risk away from financials like bond and equity futures and towards these now apparently deeper commodity markets. In a textbook example of confirmation bias, this sector rotation was misperceived, and accelerated, by some traditional money managers and discretionary hedge fund managers who interpreted the price impacts as being driven by revisions in fundamental views. They weren't. They were the result of automated portfolio construction techniques adapting to changes in market volatility.

A positive feedback spiral was kicked-off. While each CTA individually did not want to hold positions greater than 5% of the average daily volume of each market, as each CTA's allocation algo observed higher volumes, each allocated more risk, volumes went up, triggering a second round of reassessment, more allocations and so on. The real problem came just a few weeks ago in November when many of these contracts were due to be rolled. The dust hasn't even begun to settle on what happened next, but it was perhaps most succinctly described by a well-known London based trend-following pioneer (who himself claimed to have predicted and sat out these rounds of volume inflation) as 'a total sh*t show'.

## Notice periods – the elephant in the room

Matters were made worse in the days following the failed rolls when traders became aware that the redemption terms across a wide range of investment products had shrunk considerably over the last few years. While much public hoo-hah had been made of hedge funds reluctant but nonetheless gradual reduction of fees, intense competition and investor pressure had more quietly squeezed redemption notice periods. Investment products that just a few years ago would have reasonably justified notice periods of several weeks if not months, now commanded just a day.

In the days following the failed November contract rolls, some emerging market bond funds were forced to close out positions. One multi-billion dollar EM fund was said to have received redemption notices comprising some 72% of its AUM within 4 hours of the failed roll. Some of these redemption orders themselves were generated by algorithms, creating what is now being referred to as the Flash Redemption. The MSCI EM index dropped 37% in two days before market trading was suspended. Contagion spread rapidly across a range of asset classes, catalysed by the multi-strategy nature of today's trading firms.

## Déjà vu all over again

There is some hope on the horizon. While most hardened market commentators stood aghast at what appeared to be bizarre and unprecedented market behaviour, some researchers at the US Office for Financial Research recognised something familiar. These researchers had little real market experience, but this turned out to be something of a blessing. They had spent the previous few years simulating artificial market data with the use of agent-based-models, and it turned out that the recent behaviour in several markets looked remarkably like a computer simulation they had studied just a few months previously. Further digging revealed an uncanny resemblance to a chart that had appeared in a 1992 book reporting the results of a computerized simulation of a double-auction market.

The artificial computerized markets these researchers were exploring had been developed some thirty years ago by an interdisciplinary team centred around the Santa Fe Institute in the US. Although their work was largely ignored and assumed to be irrelevant to the real world at the time, in the last few months it was quickly recognised that the artificial models of the 1990s look spookily similar to today's markets.

## Science Fiction Becomes Fact

For once it seems regulators may be ahead of the private sector. In the last few years, spurred by the 2008 financial crisis, policymakers, regulators, and exchanges had begun turning to these tools for analysing issues of liquidity. Groups at the Bank of England, the US Office for Financial Research, and a network of researchers across the globe are now said to be working flat-out trying to recreate the circumstances that led to last month's market seizure.

In recent weeks the Bank of England published a paper extending their 2016 work on an agent-based model of the corporate bond market, this time applied to the copper futures market. Results demonstrated that once a small number of machine-learning competition-traders are introduced into the market, liquidity can vanish within a matter of days. In related work, economists led by a group at the New York Fed are busily introducing the role of data into their DSGE models, in much the same way that researchers quickly sought to augment the same class of models with financial features post crisis.

In a related development it is rumoured that a London based systematic fund known for its penchant towards 17th Century price data is attempting to build a proprietary data set based on the digitization of charts of simulated market data that featured in certain articles in the Journal of Economic Dynamics and Control in the mid-1990s. There are claims of the discovery of a new source of predictive signals. And Harvard based market historian Professor Philip McAvity claims to have found an uncanny resemblance between the price behaviour in lean hogs during September-November 2017 to that of the

simulation of an ant colony's foraging activity that appeared in Nature in 1996.

## Never Knowingly Undersold

Unfortunately, there appears to have been some frustration with the progress that policymakers have yet been able to make. Despite offering controversially high relative pay-scales to attract trading and risk expertise from hedge funds and trading institutions to help calibrate these new market models, it was quickly discovered that there was virtually no research expertise left within private-sector institutions. Only a handful of computation and AI graduates had entered finance in the last few years globally and most of these had been engaged in optimising office seating plans. These issues are particularly acute within quantitative systematic funds where a race to the bottom on fees has led to the total extinction of research. In one fund that claimed to be armed with over 200 PhDs, consultants discovered during an unannounced due-diligence raid that there was nothing to show for their efforts save a life-size Lego model of an elephant.

## Creating Experience

But slowly, solutions are expected to appear. As announced in the last few days on their blog, the Bank of England is said to be working with researchers at a leading social media company to create artificial researchers in an attempt to accelerate their efforts to bring real-world expertise to the model calibration process. And the Centre for the Study of Simulated Experience at the University of Arizona recently launched an online one-week intensive risk-manager training course in conjunction with a Geneva based global macro hedge-fund. This fund is one of the last remaining to have retained human risk managers. It is estimated that over 98% of hedge funds and fund-management companies outsourced their risk-management to cloud-based services during 2016.

Whether this combination of simulated experience and artificially intelligent agents will ultimately move us closer or further away from a more stable financial infrastructure remains to be seen. In the mean-time markets look towards Mr Farage and his plan to bring back liquidity "the old fashioned way" via the reintroduction of the long extinct business practice of boozy lunches. Watch this space.

**Disclaimer**

This research note is published by Neuron Advisers LLP ('Neuron'). You have been provided with this material only upon your acceptance of these Terms and Conditions herein.

Neuron Advisers LLP is authorised and regulated by the Financial Conduct Authority, with firm reference number 563919. Neuron Advisers LLP is registered with the Commodities Futures Trading Commission under NFA ID 0439462.

Investments or investment services mentioned in this research note are not available for investment and are not being marketed in any jurisdiction. The contents of this research note are not intended to be read by any persons in any jurisdiction other than the United Kingdom.

This material is presented for information purposes only. It is intended for your personal, non-commercial use. No information or opinions contained in this material constitute a solicitation or offer by Neuron to buy or sell securities or to furnish any investment advice or service. Neuron does not provide investment advice, tax advice or legal advice through this material and you agree that this material will not be used by you for such purposes.

This material is intended as a general introduction to Neuron and the Neuron research blog by which means Neuron can express ideas and opinions. The material contained herein is the sole opinion of Neuron.

The information provided in this research note is intended for institutional investors and Professional Clients and Eligible Counterparties as defined by The Financial Conduct Authority and for those who are considered as qualified eligible persons as defined by Commodities Futures Trading Commission Regulation 4.7. It is not intended for retail investors.

The contents of this research are not intended for distribution to, or use by, any individual or entity in any jurisdiction where their distribution or use would be contrary to local law or regulation or which would subject Neuron Advisers LLP to registration with the jurisdiction. You should be aware that any rules and/or regulations applicable to providing financial services (and the resultant investor protections that may be available), may not apply to persons who obtain information from the internet and its various applications, of which this material forms part.

Neuron Advisers LLP assumes no responsibility for access to this material by any person located within a country or jurisdiction where such access would be contrary to any law or regulation in that country.

We try to ensure that the information in this research note is correct, but we do not give any express or implied warranty as to its accuracy, timeliness or completeness, nor is Neuron under any obligation to update such information. Any data supplied has not been audited and is provided for information purposes only.

We are not liable for any damages (including, without limitation, damages for loss of business or loss of profits) arising in contract, tort or otherwise from the use of or inability to use this research note, or any material contained in it, or from any action or decision taken as a result of using this research or any such material.

This research note may provide links to other research and websites. We do not control the linked sites or research and we are not responsible for the contents of any linked site or research, any link in a linked site or research, or any changes or updates to such sites or research. We provide these links to you only as a convenience, and the inclusion of any link does not imply our endorsement of the site or research.

Neuron Advisers LLP is a limited liability partnership registered in England & Wales with number OC367248 and registered office address at 2nd Floor 11 Albemarle Street London W1S 4HH.

 Any questions about the contents of this material should be directed to: enquiries@neuronadvisers.com.