FINANCIAL TIMES

ft.com > comment >

May 20, 2013 6:37 pm

Financial markets are at risk of a 'big data' crash

By Maureen O'Hara and David Easley

Regulation needs to adapt to computerised details, write Maureen O'Hara and David Easley

Regulators and investors are struggling to meet the challenges posed by high-frequency trading. This ultra-fast, computerised segment of finance now accounts for most trades. HFT also contributed to the "flash crash", the sudden, vertiginous fall in the Dow Jones Industrial Average in May 2010, according to US regulators. However, the HFT of today is very different from that of three years ago. This is because of "big data".

The term describes data sets that are so large or complex (or both) that they cannot be efficiently managed with standard software. Financial markets are significant producers of big data: trades, quotes, earnings statements, consumer research reports, official statistical releases, polls, news articles, etc.

Companies that have relied on the first generation of HFT, where unsophisticated speed exploits price discrepancies, have had a tough few years. Profits from ultra-fast trading firms were 74 per cent lower in 2012 compared with 2009, according to Rosenblatt Securities. Being fast is not enough. We, along with Marcos Lopez de Prado of the Lawrence Berkeley National Laboratory, have argued that HFT companies increasingly rely on "strategic sequential trading". This consists of algorithms that analyse financial big data in an attempt to recognise the footprints left by specific market participants. For example, if a mutual fund tends to execute large orders in the first second of every minute before the market closes, algorithms able to detect that pattern will anticipate what the fund is going to do for the rest of the session, and make the same trade. The fund will keep making the trade, with higher prices, and the "algo" traders cash in.

This new form of HFT can go wrong, such as in the so-called "hash crash" of April 23 2013 – the

market drop caused by a bogus tweet about a terrorist attack on Barack Obama, sent from the Associated Press twitter feed. Unlike the crash of May 2010, this was not an incident caused by rapid sales triggering more sales. It was not a speed crash; it was a big data crash. Unless regulators understand the difference, they run the risk that new rules may address an old, expired challenge.

About two years ago, it became common for hedge funds to extract market sentiment from social media. The idea is to develop trading algorithms based on the millions of messages posted by users of Twitter, Facebook, chat rooms and blogs, and detect demand trends in relation to individual companies. However, these algorithms typically do a bad job when it comes to making guesses on small data sets. In recent months, it has become very popular to develop algorithms that fire off orders as soon as unscheduled information is published, such as natural disasters or terrorist attacks. More hash crash-type events, which are caused by a single erroneous data point, are disasters waiting to happen.

The bad news is that addressing the challenges posed by the new HFT will require understanding the mutating challenges of big data. The good news is that regulators seem to recognise the need for adaptation. This month, Scott O'Malia, Commissioner of the Commodity Futures Trading Commission, told the NYU-Poly Big Data Finance conference that "reckless behaviour" was replacing "market manipulation" as the standard for prosecuting misbehaviour. For instance, while trading on information extracted from millions of tweets is reasonable, preloading sweeping market orders as soon as an algorithm finds the words "bomb" and "White House" in the newswire is clearly reckless.

The issue at stake is, how do we make sure that participants use big data responsibly? As a Harvard Business Review article put it, big data requires big judgment. A few years ago the CFTC considered whether regulators should certify traders' algorithms. The potential for interference would be huge, not to mention the risk of intellectual property theft. A compromise may consist of market participants proposing a set of real-time indices that track "reckless" behaviour, such as adding selling pressure to a market with dwindling buyers. If a trader crosses several "recklessness" thresholds, they could be prosecuted. These indices can be changed as markets evolve and, most importantly, they could be defined by consensus among all market participants.

One solution here is to employ the resources of the US's National Laboratories. The Lawrence Berkeley National Laboratory has the supercomputing power and analysis techniques required to monitor these "recklessness" indices in real time and advise regulators of reckless market behaviour that threatens stability. While traditional circuit-breakers halt trading after a market plunge, real-time monitoring would allow for the shutting down of individual participants, preserving the market for bona fide actors.

The use of big data is transforming markets. It now needs to transform how we regulate them.

The solution to HFT issues is not less but more technology – and even bigger data.

The writers are professors at Cornell University