INTERNATIONAL CONFERENCE ON FINTECH & FINANCIAL DATA SCIENCE

3-5 September 2019, University College Dublin (UCD), Ireland

Keynote Speakers

Speaker: Nikolaus Hautsch

Email: nikolaus.hautsch@univie.ac.at

Institute & University: Department of Statistics and Operations Research, University of Vienna, Austria

Title: Limits to Arbitrage in Markets with Stochastic Settlement Latency**:**

Abstract:

Distributed ledger technologies rely on consensus protocols confronting traders with random waiting times until the transfer of ownership is accomplished. This time-consuming settlement process exposes arbitrageurs to price risk and imposes limits to arbitrage. We derive theoretical arbitrage boundaries under general assumptions and show that they increase with expected latency, latency uncertainty, spot volatility, and risk aversion. Using high-frequency data from the Bitcoin network, we estimate arbitrage boundaries due to settlement latency of on average 124 basis points, covering 88% of the observed cross-exchange price differences. We document cross-exchange flows chasing arbitrage opportunities only if we account for transaction cost and settlement latency. Settlement through decentralized systems thus induces non-trivial frictions affecting market efficiency and price formation.

Short Bio:

Nikolaus Hautsch is Professor of Finance and Statistics at the University of Vienna. He earned his Ph.D. in econometrics in 2003 from the University of Konstanz. From 2004 to 2007 he joined the Department of Economics of the University of Copenhagen as Assistant Professor and Associate Professor. Until 2013 he held the Chair of Econometrics at Humboldt University Berlin and was director of the Berlin Doctoral Program in Economics and Management Science. He is elected fellow of the Society for Financial Econometrics, research fellow of the Center for Financial Studies (CFS) Frankfurt and a staff member of the Vienna Graduate School of Finance. Hautsch had visiting positions at the University of Technology, Sydney, the University of Melbourne, the Université Catholique de Louvain, the University of Cambridge and Duke University. His research focuses on the econometrics of high-frequency financial data, market microstructure analysis, the modelling of volatility and liquidity, systemic risk and information processing on financial markets. He publishes in leading journals in the area of finance, econometrics and statistics. Currently, he serves as associate editor of the Journal of Business and Economic Statistics, the Journal of Applied Econometrics, the Journal of Financial Econometrics and the International Journal of Forecasting, among others.

Speaker: Ser-huang Poon

Email: ser-huang.poon@manchester.ac.uk

Institute & University: Alliance Manchester Business School, University of Manchester, UK

Title: Responsible NLP for Modern Slavery Intelligence

Authors: Mark Andreev and Irina Goloshchapova (Lomonosov Moscow State University); Matthew Pritchard and Ser-Huang Poon (Alliance Manchester Business School)

Abstract:

With the Modern Slavery Act 2015, the UK became the first country to require organisations to publicly

report on the steps they are taking to prevent modern slavery in their operations and supply chains. Current Home Office's estimation suggests 17,000 UK-based organisations will be in scope. The home office is tasked with carrying out an audit of compliance, further to this audit publicly name non-compliant organisations. Auditing textual disclosure of 17,000 organisations is not a trivial exercise; our NLP work aims to facilitate the human effort with machine guided assessment indicators. In particular, we use NLP analysis and a machine-human guided semi-supervised approach to map modern slavery incidents around the world from historical and current news coverage of modern slavery practices. A pilot study using TF-IDF aided Logistic Regression and Bloomberg child and forced labour news from 2014 to 2019, shows we were able to predict an actual incident with 78% precision and a false incident with 79% of precision. We then use semi supervise topic modelling (LDA) and Wikipedia supported Named-Entity Recognition to start mapping modern slavery types and distribution around the world. This intelligence is then feedback to supply chain surveillance as intended under the Modern Slaver Act.

Short Bio:

Dr Ser-Huang Poon is a Professor at the University of Manchester, a Turing Fellow at the Turing Institute for Data Science & Artificial Intelligence, and the founder and trustee of the charity, EnduringNet (https://enduringnet.org/), a consortium of academia, researchers, blockchain developers, IT and business professionals, and passionate advocates – to leverage distributed ledger technologies (DLT) and artificial intelligence (AI) in humanitarian work. Her work was cited on the Nobel website as a reference reading in volatility. Dr Poon received two best paper awards, and she has co-authored a commissioned report for the Her Majesty Treasur Foresight Program on the impact of high frequency trading. In 2018, she submitted evidence for the Independent Review of the MSA, Section 54 -Transparency in Supply Chains. Her current research interests evolve around Blockchain, FinTech, AI, BigData and Corporate Social Responsibility.

Speaker: Hugh Christensen

Email: hughchristensen@bmlltech.com

Institute & University: BMLL Technologies

Title: Enabling Value Extraction from Limit Order Book Data

Abstract:

L3 limit order book (LOB) data contains 30% more price discovery relevant information than L1 data. The problem is this data is difficult to work with and the information is hard to extract. While a small handful of very well resourced leading market participants have capitalized on being able to use this data over the last decade, most major financial institutions are still not able to.

Some believe that L3 information is not relevant to those trading through brokers, nor relevant to those trading at medium to low frequencies. To which we would make three observations. Firstly, while it is true that L3 data does hold information on short-term patterns, it is only by down-sampling this data that lower frequency data is obtained. Pre-processing with a zero-one filter is a crude tool, throwing away information whether it is relevant or not. The statistically correct approach is to allow the model to decide how to handle the data. Secondly, the "leading market participants" referred to above generally ply their trade through what many consider to be 'predatory' trading. This can mean providing liquidity when it suits, but as soon as a natural order is detected in the L3 data, leaning against that order and taking liquidity – the end result being a cost to the provider of the natural order. Only by understanding the statistical dynamics of parent-child submission through using L3 data can such predatory approaches be defeated. Thirdly is the ability to accurately simulate the market using agent based models (ABMs), trained using L3 data. From trained ABMs synthetic market data can be generated. The ability to simulate medium-low frequency trading strategies not just against a one-time realization of market data, but across many realizations enables a step change in statistical significance of predictor design.

The workflow of using L3 data is as follows: Data is collected by performing packet capture at the colo

with potentially every pcap location recording data from every matching engine. This captured data is then centralized and subject to a process of curation (eg book building, ticker mapping), normalization (eg transform to UTC, map fields to API dictionaries) and consolidation (eg enable my view of the European Consolidated Tape). An additional high-value step is combing the L3 public data with private order flow data to generate L4 data – enabling identification of beneficiaries orders in the book, along with other information such as order types and max show values. This data pipeline will include metadata management and potentially derived data management (eg generate intraday volume curves). Once the data is present it then needs to be combined with cheap compute at scale. This is either by an in-house farm or using the cloud, using either map-reduce or non-map-reduce systems. The workflow needs to allow for the range of L3 data use cases and where and how to implement fine grained and coarse grained parallelism to ensure sufficient speed. Finally, the workflow needs access to both open source analytics (eg pandas, tensorflow) and closed source analytics (eg MOSEK). The end user needs to be able to make any arbitrary calculation on any amount of L3 data (eg Russell 3000 for the last three years) and have results returned to them in an appropriate amount of time. At BMLL we see our value as being at the interfaces of this value stack, with the interfaces being represented as APIs.

A recent industry survey of quants found they felt they spent over 80% of their time performing menial duties around data and systems, as opposed to performing their value-add. Another recent survey of employers felt that they were severely under-resourced in the quant and engineering fields. In summary, managed services are powerful as they help solve for both these pain points. Managed services enable human resources to focus on their value add, while also speeding up the OODA loop of complex problem solving.

Short Bio: https://uk.linkedin.com/in/hughchristensen

Speaker: Mark Andrew Chen

Email: machen@gsu.edu

Institute & University: Georgia State University, U.S.A.

Title:

Abstract:

Short Bio:
https://robinson.gsu.edu/profile/mark-a-chen/

Speaker: Jyrki Piilo

Email: jyrki.piilo@utu.fi

Institute & University: Turku Centre for Quantum Physics, Department of Physics and Astronomy, University of Turku, Finland

Title: Networks of investors in financial markets

Abstract:

During the last 20 years the use of network methods have become common in various fields of science ranging, e.g., from biology to technological and social systems [1]. The abundance of data financial markets produce opens also intriguing possibilities to use complex networks to study various financial systems [2]. We have both developed methods for network theory [3,4] and applied them, during the last ten years, to trading data in financial markets [5-8].

Empirical data is often noisy and heterogenous — and it is not obvious how to extract the structure and properties of a given system. For this purpose we have developed a method of statistically validated networks (SVNs) allowing to construct the most informative structure of the network and its

characterising properties [3-4].

By having access to unique 20 year database containing trading histories of all investors trading Finnish stocks, we have constructed a network of investors with subsequent observation of their clustering based on their actions in stock markets [5]. The methods and data also allow to study the long-term time evolution of the investor networks [8]. Here, our empirical observations show the presence of an ecology of groups of investors characterised by different attributes and by various investment styles over many years. Moreover, our findings suggest that groups of traders are always competing, adopting, using and eventually discarding new investment strategies. This adaptation process is observed over a multiplicity of time scales, and is compatible with several conclusions of behavioral finance and with the assumptions of the so-called adaptive market hypothesis [9]. Time allowing, we may also describe a networked structure of HFT on the level of market members.

[1] M. E. J. Newman, Networks An Introduction, Oxford University Press, New York, 2010.

[2] Focus issue: Complex networks in finance, Nature Physics 9 (2013) 119–197.

[3] M. Tumminello, S. Micciche, F. Lillo, J. Piilo, and R.N. Mantegna, PLoS ONE 6, e17994 (2011).

[4] M. Tumminello, , S. Miccichè, F. Lillo, , J. Varho, J. Piilo, R. N. Mantegna, Journal of Statistical Mechanics: Theory and Experiment (2011) P01019.

[5] M. Tumminello, F. Lillo, J. Piilo, and R. N. Mantegna, New Journal of Physics 14, 013041 (2012).

[6] F. Lillo, S. Miccichè, M. Tumminello, J. Piilo, R. N. Mantegna, Quantitative Finance 15, (2015) 213–229.

[7] F. Musciotto, L. Marotta, S. Miccichè, J. Piilo, and R. N. Mantegna, Chaos, Solitons and Fractals, 88, 267 (2016).

[8] F. Musciotto, L. Marotta, J. Piilo, and R. N. Mantegna, "Long-term ecology of investors in a financial market", Palgrave Communications 4, 92 (2018).

[9] A. W. Lo, "The adaptive market hypothesis", J. Portf. Manag. 30, 15 (2004).

Short Bio: http://users.utu.fi/jyrpii/