August 2022, No 6

# Quantitative Finance Research Newsletter
# Oxford-Man Institute
# OMIReNew




*The Oxford-Man Institute of Quantitative Finance is a world-leading centre for interdisciplinary research in financial markets. This newsletter is an eclectic summary that covers the most recent research of relevance to quantitative finance, from a wide range of disciplines in academia and industry.*

***This issue**: Machine Learning for Finance, Graph-based Learning, Derivatives & Volatility, Quantitative Finance & Economics, Natural Language Processing, and Digital Assets.*

## Highlights

Noteworthy works in this month's edition include: a work from OMI, introducing algorithmic learning equations showing that algorithms can learn, and coordinate, a reward-punishment mechanism to sustain tacit collusion; a multi-task learning approach from Nanyang Technological University for investments by routing uncertainty aware trading experts; a work from J.P. Morgan applying continuous reinforcement learning to hedge portfolios, and a work from Aarhus University predicting trading behaviour of socially connected investors.

## Machine Learning for Finance

### The Algorithmic Learning Equations: Evolving Strategies in Dynamic Games

A. Cartea, P. Chang, J. Penalva, and H. Waldon
Uni. of Oxford & Universidad Carlos III

This work introduces algorithmic learning equations (ALEs), a set of ordinary differential equations which characterises the finite-time and asymptotic behaviour of the stochastic interaction between state-dependent learning algorithms in dynamic games. The ALEs, new to the literature, admit a rigorous characterization of the dynamics of a wide class of general learning algorithms, including well-known algorithms such as asynchronous, state-dependent Q-learning. Their framework includes a variety of information and memory structures, including noisy, perfect, private, and public monitoring. They demonstrate their methodology in a repeated 2x2 prisoner's dilemma game with perfect monitoring. Their analysis of the ALEs shows that algorithms can learn a reward-punishment mechanism to sustain tacit collusion and can also learn to coordinate in cycles of mutual collusion and mutual punishment.

### Quantitative Stock Investment by Routing Uncertainty-Aware Trading Experts: A Multi-Task Learning Approach

S. Sun, R. Wang, and B. An
Nanyang Technological University

Quantitative investment is a fundamental financial task that highly relies on accurate stock prediction and profitable investment decision making. Despite recent advances in deep learning (DL) have shown stellar performance on capturing trading opportunities in the stochastic stock market, they observe that the performance of existing DL methods is sensitive to random seeds and network initialization. To design more profitable DL methods, they analyse this phenomenon and find two major limitations of existing works. First, there is a noticeable gap between accurate financial predictions and profitable investment strategies. Second, investment decisions are made based on only one individual predictor without consideration of model uncertainty, which is inconsistent with the workflow in real-world trading firms. To tackle these two limitations, they first reformulate quantitative investment as a multi-task learning problem. Later on, they propose AlphaMix, a novel two-stage mixture-ofexperts (MoE) framework for quantitative investment to mimic the efficient bottom-

up trading strategy design workflow of successful trading firms. In Stage one, multiple independent trading experts are jointly optimized with an individual uncertainty-aware loss function. In Stage two, they train neural routers (corresponding to the role of a portfolio manager) to dynamically deploy these experts on an as-needed basis. AlphaMix is also a universal framework that is applicable to various backbone network architectures with consistent performance gains. Through extensive experiments on long-term real-world data spanning over five years on two of the most influential financial markets (US and China), they demonstrate that AlphaMix significantly outperforms many state-of-the-art baselines in terms of four financial criteria. Furthermore, they analyse the contributions of each component in AlphaMix with a series of exploratory and ablative studies.

### Double Deep Q-Learning for Optimal Execution

B. Ning, F. Lin, and S. Jaimungal
University of Toronto

Optimal trade execution is an important problem faced by essentially all traders. Much research into optimal execution uses stringent model assumptions and applies continuous time stochastic control to solve them. Here, they instead take a model free approach and develop a variation of Deep Q-Learning to estimate the optimal actions of a trader. The model is a fully connected Neural Network trained using Experience Replay and Double DQN with input features given by the current state of the limit order book, other trading signals, and available execution actions, while the output is the Q-value function estimating the future rewards under an arbitrary action. They apply their model to nine different stocks and find that it outperforms the standard benchmark approach on most stocks using the measures of (i) mean and median out-performance, (ii) probability of out-performance, and (iii) gain-loss ratios.

### Deep Reinforcement Learning for Market Making Under a Hawkes Process-Based Limit Order Book Model

B. Gasperov, and Z. Kostanjcar
University of Zagreb

The stochastic control problem of optimal market making is among the central problems in quantitative finance. In this paper, a deep reinforcement learning-based controller is trained on a weakly consistent, multivariate Hawkes process-based limit order book simulator to obtain market making controls. The proposed approach leverages the advantages of Monte Carlo backtesting and contributes to the line of research on market making under weakly consistent limit order book models. The ensuing deep reinforcement learning controller is compared to multiple market making benchmarks, with the results indicating its superior performance with respect to various risk-reward metrics, even under significant transaction costs.

### Autoencoding Conditional GAN for Portfolio Allocation Diversification

J. Lu, and S. Yi
Unaffiliated

Over the decades, the Markowitz framework has been used extensively in portfolio analysis though it puts too much emphasis on the analysis of the market uncertainty rather than on the trend prediction. While generative adversarial network (GAN) and conditional GAN (CGAN) have been explored to generate financial time series and extract features that can help portfolio analysis. The limitation of the CGAN framework stands in putting too much emphasis on generating series rather than keeping features that can help this generator. In this paper, they introduce an autoencoding CGAN (ACGAN) based on deep generative models that learns the internal trend of historical data while modeling market uncertainty and future trends. They evaluate the model on several real-world datasets from both the US and Europe markets, and show that the proposed ACGAN model leads to better portfolio allocation and generates series that are closer to true data compared to the existing Markowitz and CGAN approaches.

## Graph-based Learning

### HATR-I: Hierarchical Adaptive Temporal Relational Interaction for Stock Trend Prediction

H. Wang, T. Wang, S. Li, and S. Guan
Peking University

Stock trend prediction is a hot issue in the Fintech field. Effective stock profiling is challenging due to highly non-stationary dynamics and complex interplays. Existing methods usually regard each stock independently or detect simplistic homogeneous structures. Practically, stock correlation originates from diverse aspects and underlying relationship signals are implicit in comprehensive graphs. Besides, RNNs are extensively used to simulate stock volatility while inadequate in capturing fine-granular patterns across local time snippets. To this end, in this paper they propose HATR-I, a Hierarchical Adaptive Temporal-Relational Interaction model to characterise and predict stock evolutions. Specifically, they grasp short- and long-term transition regularities of stock dynamics based on cascaded dilated convolutions and gating paths. By formulating different views of domain adjacency graphs into a unified multiplex network with edge attributes, they inject node- and semantic-level dual attention to refine the propagation of inter-stock

collaborative information. Particularly, the stock pair matching is proceeding along each time-stage rather than until final compressed representations, meanwhile significant feature points and scales are identified considering the effect of time attenuation. Finally, they deduce latent shared clusters as global regularization to optimise the stock representations. Experiments on three real-world stock market datasets demonstrate the effectiveness of their proposed model.

### Asset Pricing via Deep Graph Learning to Incorporate Heterogeneous Predictors

J. Huang, R. Xing, and Q. Li
Southwestern Uni. of Finance and Economics

Stock trend prediction is a hot issue in the Fintech field. Effective stock profiling is challenging due to highly non-stationary dynamics and complex interplays. Existing methods usually regard each stock independently or detect simplistic homogeneous structures. Practically, stock correlation originates from diverse aspects and underlying relationship signals are implicit in comprehensive graphs. Besides, RNNs are extensively used to simulate stock volatility while inadequate in capturing fine-granular patterns across local time snippets. To this end, they propose HATR-I, a Hierarchical Adaptive Temporal-Relational Interaction model to characterise and predict stock evolutions. Specifically, they grasp short- and long-term transition regularities of stock dynamics based on cascaded dilated convolutions and gating paths. By formulating different views of domain adjacency graphs into a unified multiplex network with edge attributes, they inject node- and semantic-level dual attention to refine the propagation of inter-stock collaborative information. Particularly, the stock pair matching is proceeding along each time-stage rather than until final compressed representations, meanwhile significant feature points and scales are identified considering the effect of time attenuation. Finally, they deduce latent shared clusters as global regularization to optimise the stock representations. Experiments on three real-world stock market datasets demonstrate the effectiveness of their proposed model.

### Predicting the Trading Behaviour of Socially Connected Investors.

K. Baltakys, M. Baltakiene, N. Heidari, A. Iosifidis, and J. Kanniainen
Aarhus Uni. & Tampere Uni.

This work finds that investors' future trading decisions are driven by the patterns of their social neighbourhood and the trading activity therein. Moreover, they provide evidence that investors weigh their social connections differently in terms of information transfer. Methodologically, they tackle the complex, cyclical patterns of investor social networks by graph neural networks, which allow us to propose a sophisticated way to predict the behavior of investors with data on their social connections. Their analysis is based on the unique data on observed social links through director (insider) positions on the same companies as well as links to family members, together with full investor-level market-wise transaction data. They make the anonymized data set open to the public.

### Subsequence-based Graph Routing Network for Capturing Multiple Risk Propagation Processes

R. Cheng, and Q. Li
Southwestern Uni. of Finance and Economics

In finance, the risk of an entity depends not only on its historical information but also on the risk propagated by its related peers. Pilot studies rely on Graph Neural Networks (GNNs) to model this risk propagation, where each entity is treated as a node and represented by its time-series information. However, conventional GNNs are constrained by their unified messaging mechanism with an assumption that the risk of a given entity only propagates to its related peers with the same time lag and has the same effect, which is against the ground truth. In this study, they propose the subsequencebased graph routing network (S-GRN) for capturing the variant risk propagation processes among time-series represented entities. In S-GRN, the messaging mechanism between each node pair is dynamically and independently selected from multiple messaging mechanisms based on the dependencies of variant subsequence patterns. The SGRN is extensively evaluated on two synthetic tasks and three real-world datasets and demonstrates state-of-the-art performance.

## Derivatives & Volatility

### Solving Barrier Options Under Stochastic Volatility Using Deep Learning

W. Fu and A. Hirsa
Columbia University

This work develops an unsupervised deep learning method to solve the barrier options under the Bergomi model. The neural networks serve as the approximate option surfaces and are trained to satisfy the PDE as well as the boundary conditions. Two singular terms are added to the neural networks to deal with the non-smooth and discontinuous payoff at the strike and barrier levels so that the neural networks can replicate the asymptotic behaviors of barrier options at short maturities. After that, vanilla options and barrier options are priced in a single framework. Also, neural networks are employed to deal with the high dimensionality of the

function input in the Bergomi model. Once trained, the neural network solution yields fast and accurate option values.

## Exchange Option Pricing Under Variance Gamma-like Models

M. Gardini, and P. Sabino
Uni. of Genoa, Uni. of Helsinki & E.ON

In this article they focus on the pricing of exchange options when the dynamic of logprices follows either the well-known variance gamma or the recent variance gamma++ process. In particular, for the former model they can derive a Margrabe's type formula whereas, for the latter one they can write an "integral free" formula. Furthermore, they show how to construct a general multidimensional version of the variance gamma++ processes preserving both the mathematical and numerical tractability. Finally, they apply the derived models to German and French energy power markets: they calibrate their parameters using real market data and they accordingly evaluate exchange options with the derived closed formulas, Fourier-based methods and Monte Carlo techniques.

## Volatility Based Kernels and Moving Average Means for Accurate Forecasting with Gaussian Processes

G. Benton, W. Maddox, and A. Wilson
New York University

A broad class of stochastic volatility models are defined by systems of stochastic differential equations. While these models have seen widespread success in domains such as finance and statistical climatology, they typically lack an ability to condition on historical data to produce a true posterior distribution. To address this fundamental limitation, they show how to re-cast a class of stochastic volatility models as a hierarchical Gaussian process (GP) model with specialized covariance functions. This GP model retains the inductive biases of the stochastic volatility model while providing the posterior predictive distribution given by GP inference. Within this framework, they take inspiration from well studied domains to introduce a new class of models, Volt and Magpie, that significantly outperform baselines in stock and wind speed forecasting, and naturally extend to the multitask setting.

## Volatility Based Kernels and Moving Average Means for Accurate Forecasting with Gaussian Processes

G. Benton, W. Maddox, and A. Wilson
New York University

A broad class of stochastic volatility models are defined by systems of stochastic differential equations. While these models have seen widespread success in domains such as finance and statistical climatology, they typically lack an ability to condition on historical data to produce a true posterior distribution. To address this fundamental limitation, they show how to re-cast a class of stochastic volatility models as a hierarchical Gaussian process (GP) model with specialized covariance functions. This GP model retains the inductive biases of the stochastic volatility model while providing the posterior predictive distribution given by GP inference. Within this framework, they take inspiration from well studied domains to introduce a new class of models, Volt and Magpie, that significantly outperform baselines in stock and wind speed forecasting, and naturally extend to the multitask setting.

## Continuous Reinforcement Learning for Hedging of General Portfolios across Multiple Risk Aversions

P. Murray, B. Wood, H. Buehler, M. Wiese, and M. Pakkanen
J.P. Morgan, TU Munich, TU Kaiserslautern & Imperial College

A broad class of stochastic volatility models are defined by systems of stochastic differential equations. While these models have seen widespread success in domains such as finance and statistical climatology, they typically lack an ability to condition on historical data to produce a true posterior distribution. To address this fundamental limitation, they show how to re-cast a class of stochastic volatility models as a hierarchical Gaussian process (GP) model with specialized covariance functions. This GP model retains the inductive biases of the stochastic volatility model while providing the posterior predictive distribution given by GP inference. Within this framework, they take inspiration from well studied domains to introduce a new class of models, Volt and Magpie, that significantly outperform baselines in stock and wind speed forecasting, and naturally extend to the multitask setting.

# Quantitative Finance & Economics

## Portfolio Construction as Linearly Constrained Separable Optimisation

N. Moehle, J. Gindi, S. Boyd, and M. Kochenderfer
Stanford University

Mean-variance portfolio optimization problems often involve separable nonconvex terms, including penalties on capital gains, integer share constraints, and minimum

position and trade sizes. They propose a heuristic algorithm for such problems based on the alternating direction method of multipliers (ADMM). This method allows for solve times in tens to hundreds of milliseconds with around 1000 securities and 100 risk factors. They also obtain a bound on the achievable performance. Their heuristic and bound are both derived from similar results for other optimization problems with a separable objective and affine equality constraints. They discuss a concrete implementation in the case where the separable terms in the objective are piecewise quadratic, and they empirically demonstrate its effectiveness for tax-aware portfolio construction.

## A Horse Race of Monetary Policy Regimes: An Experimental Investigation

O. Kostyshyna, L. Petersen, and J. Yang
Bank of Canada

This work provides a comprehensive assessment of five monetary policy regimes—inflation targeting (IT), dual mandate (DM), average inflation targeting under 4-period (AIT-4) and 10-period (AIT-10) horizons, price level targeting (PLT), and nominal GDP level targeting (NGDP)—in a unified experimental framework. They study how participants can understand different regimes and form expectations during periods of economic stability and during a demand-driven recession that temporarily brings the economy to the ELB. Their results suggest a distinct ranking of policy regimes in terms of their ability to achieve macroeconomic stability. DM and IT are the most stabilizing regimes, followed by AIT and then level-targeting regimes PLT and NGDP. AIT with a shorter horizon performs better than AIT with a longer horizon. Monetary policy regimes that are framed around the inflation rate (e.g., AIT-10) are found to deliver significantly more stable economic outcomes than those that target price levels (PLT). Participants have greater difficulty understanding regimes that are more history-dependent and forecasting in the rationally expected direction. They instead rely on trend-chasing heuristics to form their expectations. Trend-chasing forecasting is more destabilizing for regimes with more history dependence. Participants also "need to see it to believe it." PLT and NGDP initially have mixed success at achieving their targets, and these regimes do not gain credibility before the economy enters into the ELB, where credibility is needed more than ever.

## Distributionally Robust Inventory Management with Advance Purchase Contracts

Y. Xue, Y. Li, and N. Rujeerapaiboon
National Uni. of Singapore, and Southeast Uni.

They propose a distributionally robust inventory model for finding an optimal ordering policy that attains the minimum worst-case expected total cost. In a classical stochastic setting, this problem is typically addressed by dynamic programming and is solved by the famous base-stock policy. This approach however crucially relies on two controversial assumptions: the demands are serially independent and the demand distribution is perfectly known. Aiming to address these issues, they adopt a mean-variance ambiguity set that imposes neither the shape of each marginal demand distribution nor their independence structure, and they focus on the case of advance purchase agreements which are prevalent in the robust inventory literature and have drawn renewed attention because of the Covid-19 vaccine procurement. The proposed distributionally robust inventory model provably reduces to a finite conic optimization problem with however an exponential number of constraints. To gain tractability and to err on the safe side, they propose two conservative approximations. The first approximation is obtained by recognizing the problem as an artificial two-stage robust optimization problem and then by restricting each adaptive decision to a linear decision rule. The second approximation, on the other hand, is obtained by a constraint partitioning and by upper bounding each resultant maximum sum with a sum of maxima. They then present a progressive approximation based on a scenario reduction technique to gauge the quality of the proposed conservative approximations. They prove that this progressive approximation is exact when the inventory problem consists of two periods, and besides they use it to show that their conservative solutions are still close to being optimal when the planning horizon is longer. All of their exact and approximate inventory models are expressed as standard conic programs which allow for the incorporation of additional distributional information. The extensions are readily obtained by deriving a new cone that corresponds to the restricted ambiguity set and embedding it in the original problems. They analytically derive the worst-case demand distribution from the mean-variance ambiguity set and numerically use it to show that their robust inventory policy is more resilient to the misspecification of the demand distribution than the state-of-the-art non-robust policies.

## Monetary Policy & Anchored Expectations: An Endogenous Gain Learning Model

L. Gati
European Central Bank

This paper analyses monetary policy in a model with a potential unanchoring of inflation expectations. The degree of unanchoring is given by how sensitively the public's long-run inflation expectations respond to inflation surprises. The author finds that optimal policy moves the interest rate aggressively when expectations

unanchor, allowing the central bank to accommodate inflation fluctuations when expectations are well-anchored. Furthermore, the author estimate the model-implied relationship that determines the extent of unanchoring. The data suggest that the expectations process is nonlinear and asymmetric: expectations respond more sensitively to large or downside surprises than to smaller or upside ones.

### Application of Hawkes volatility in the observation of filtered high-frequency price process in tick structures

K. Lee
Yeungnam University

The Hawkes model is suitable for describing self and mutually exciting random events. In addition, the exponential decay in the Hawkes process allows us to calculate the moment properties in the model. However, due to the complexity of the model and formula, few studies have been conducted on the performance of Hawkes volatility. In this study, they derived a variance formula that is directly applicable under the general settings of both unmarked and marked Hawkes models for tick-level price dynamics. In the marked model, the linear impact function and possible dependency between the marks and underlying processes are considered. The Hawkes volatility is applied to the mid-price process filtered at 0.1-second intervals to show reliable results; furthermore, intraday estimation is expected to have high utilization in real-time risk management. They also note the increasing predictive power of intraday Hawkes volatility over time and examine the relationship between futures and stock volatilities.

## Natural Language Processing

### Learning Mutual Fund Categorization using Natural Language Processing

D. Vamvourellis, M. Toth, D. Desai, D. Mehta, and S. Pasquali

BlackRock

Categorization of mutual funds or Exchange-Traded-funds (ETFs) have long served the financial analysts to perform peer analysis for various purposes starting from competitor analysis, to quantifying portfolio diversification. The categorization methodology usually relies on fund composition data in the structured format extracted from the Form N-1A. Here, they initiate a study to learn the categorization system directly from the unstructured data as depicted in the forms using natural language processing (NLP). Positing as a multi-class classification problem with the input data being only the investment strategy description as reported in the form and the target variable being the Lipper Global categories, and using various NLP models, they show that the categorization system can indeed be learned with high accuracy. They discuss implications and applications of their findings as well as limitations of existing pre-trained architectures in applying them to learn fund categorization.

## Digital Assets

### Trading Under The Proof-of-Stake Protocol - A Continuous-time Control Approach

W. Tang, and D. Yao
Columbia University

They develop a continuous-time control approach to optimal trading in a Proof-of-Stake (PoS) blockchain, formulated as a consumption-investment problem that aims to strike the optimal balance between a participant's (or agent's) utility from holding/trading stakes and utility from consumption. They present solutions via dynamic programming and the Hamilton-Jacobi-Bellman (HJB) equations. When the utility functions are linear or convex, they derive close-form solutions and show that the bang-bang strategy is optimal (i.e., always buy or sell at full capacity). Furthermore, they bring out the explicit connection between the rate of return in trading/holding stakes and the participant's risk-adjusted valuation of the stakes. In particular, they show when a participant is risk-neutral or riskseeking, corresponding to the risk-adjusted valuation being a martingale or a sub-martingale, the optimal strategy must be to either buy all the time, sell all the time, or first buy then sell, and with both buying and selling executed at full capacity. They also propose a risk-control version of the consumption-investment problem; and for a special case, the "stake-parity" problem, they show a mean-reverting strategy is optimal.

### Slippage in AMM Markets

I. Aldrige
University of Cambridge

The crypto market is growing at an unprecedented rate, but it can be difficult to make money in this market if you don't understand how it works. Most people are still using traditional methods to trade in the crypto market, which can lead to losses due to significantly adverse pricing, known as slippage. In this paper, they 1) explain the key differences between Automated Market Making (AMM) used in crypto trading and traditional limit-order book-based market making, 2) show crypto price formation step-by-step, and 3) derive a closed-form solution for ex-ante crypto slippage estimation that can be used to extract better profits from crypto trading.

## Fragmentation, Price Formation and Cross-Impact in Bitcoin Markets

J. Albers, M. Cucuringu, S. Howison, and A. Shestopaloff

University of Oxford

In the light of micro-scale inefficiencies due to the highly fragmented bitcoin trading landscape, they use a granular data set comprising orderbook and trades data from the most liquid bitcoin markets, to understand the price formation process at sub-1-second time scales. To this end, they construct a set of features that encapsulate relevant microstructural information over short lookback windows. These features are subsequently leveraged, first to generate a leader–lagger network that quantifies how markets impact one another, and then to train linear models capable of explaining between 10% and 37% of total variation in 500 ms future returns (depending on which market is the prediction target). The results are then compared with those of various PnL calculations that take trading realities, such as transaction costs, into account. The PnL calculations are based on natural taker strategies (meaning they employ market orders) associated with each model. Their findings emphasise the role of a market's fee regime in determining both its propensity to lead or lag, and the profitability of their taker strategy. They further derive a natural maker strategy (using only passive limit orders) which, due to the difficulties associated with backtesting maker strategies, they test in a real-world live trading experiment, in which they turned over 1.5 M USD in notional volume. Lending additional confidence to their models, and by extension to the features they are based on, the results indicate a significant improvement over a naive benchmark strategy, which they also deploy in a live trading environment with real capital, for the sake of comparison.

## A Luna-tic Stablecoin Crash

H. Uhlig

University of Chicago

After remaining close to 1 US Dollar since its inception in November 2020, the algorithmic stablecoin UST crashed in the two weeks of May 9th to May 15th, 2022, leading to a price collapse of the underlying LUNA token and the erasure of more than 50 Billion U.S. Dollar or 90% in market value. The author provides a novel theory to account for these phenomena and use it to shed light on the data. The author breaks new ground by showing how crashes unfold gradually, and by introducing the method of quantitative interpretation. To obtain a gradual unfolding of the crash, the author allows for the possibility that the market might return to normal at any moment. Suspension of convertibility happens, once the price has fallen sufficiently far. Agents price LUNA, taking into account these probabilities as well as the ongoing inflow from burning UST coins. Agents sell their UST coins, when the probability of an eventual suspension of convertibility exceeds some agent-specific threshold. The implications of the theory are highlighted in an analytically tractable example. The theory is then used as a guide to examine and interpret the data, using bi-hourly observations. The author uses the observed data to quantify theory variables, and use them in turn to interpret the data. The work finds that the majority of the UST coin holders waited until the probability of suspension was rather high, before deciding to burn their holdings.

# Acknowledgement

This newsletter was curated by Utkarsh Sharma, D.Phil. Candidate, and Prof. Álvaro Cartea, Director, at the Oxford-Man Institute, with paper suggestions from fellow students at the institute - Dragos Gorduza, Marcello Monga, and Patrick Chang.