OMI Research Newsletter – July 2025



The Oxford-Man Institute of Quantitative Finance (OMI) is a world-leading centre for interdisciplinary research in financial markets. The OMI addresses fundamental problems in quantitative finance with a strong focus on machine learning and data driven models.

*The **OMI Research Newsletter** is a curated summary that covers the most recent research, insights, and tools related to financial decision making.*

Announcements

The OMI is seeking to appoint one full-time Postdoctoral Research Assistant in Blockchain Economics and Decentralised Finance. Please find more details here.

Highlights

Key contributions in this month's edition: Duran-Martin et al. (2025) propose a scalable algorithm for online learning and generalised Bayesian inference of neural network parameters, which combines the strengths of frequentist Kalman-style updates and Bayesian filtering. Vuletić and Cucuringu (2025) introduce GraFiN-Gen, a graph-based ensemble framework for multi-asset forecasting based on profit-driven connections. Andersen, Ding and Todorov (2025) develop a nonparametric high-frequency estimator for the tail shape of cross-sectional asset return distributions, and show that the tail behaviour differs in the presence or absence of systematic jump events. Capponi, Cartea and Drissi (2025) develop a theoretical framework to illustrate how block time and priority-fee auctions in decentralised exchanges affect price efficiency, market frictions, and liquidity. Nagel (2025) and Kelly and Malamud (2025) have further discussions about the "virtue of complexity" in return prediction, with fresh theoretical insights and new empirical evidence. Chen et al. (2025) identify a parsimonious nine-factor model for hedge fund returns, and demonstrate its superior performance compared to existing models. Baggiani, Herdegen and Sánchez-Betancourt (2025) derive an optimal dynamic fee schedule for constant-function automated market makers, and illustrate its economic intuition for the arbitrageurs/noise traders trade-off.

Statistics, Econometrics & Machine Learning

Scalable Generalized Bayesian Online Neural Network Training for Sequential Decision Making

G. DURAN-MARTIN, L. SÁNCHEZ-BETANCOURT, Á. CARTEA, K. MURPHY

UNIVERSITY OF OXFORD, GOOGLE DEEPMIND

We introduce scalable algorithms for online learning and generalized Bayesian inference of neural network parameters, designed for sequential decision making tasks. Our methods combine the strengths of frequentist and Bayesian filtering, which include fast low-rank updates via a block-diagonal approximation of the parameter error covariance, and a well-defined posterior predictive distribution that we use for decision making. More precisely, our main method updates a low-rank error covariance for the hidden layers parameters, and a full-rank error covariance for the final layer parameters. Although this characterizes an improper posterior, we show that the resulting posterior predictive distribution is well-defined. Our methods update all network parameters online, with no need for replay buffers or offline retraining. We show, empirically, that our methods achieve a competitive tradeoff between speed and accuracy on (non-stationary) contextual bandit problems and Bayesian optimization problems.

GraFiN-Gen: Graph-Based Ensemble Generative Modelling for Multi-Asset Forecasting

M. VULETIĆ, M. CUCURINGU

UNIVERSITY OF OXFORD, UNIVERSITY OF CALIFORNIA LOS ANGELES

We extend conditional generative models to the multi-asset setting by leveraging cross-asset relationships through a graph-based probabilistic ensemble framework. Rather than combining pointwise forecasts, our method ensembles full conditional return distributions. The graph captures the transferability of predictive information across assets, with edge weights learned via an economic profit-maximisation objective reformulated as a LASSO regression. This induces a sparse, interpretable, and computationally efficient calculation of weights, accounting for the correlations between forecasts. We apply our methodology to Fin-GAN, a custom conditional GAN trained to forecast the directionality of asset moves, and demonstrate that the LASSO-induced graph outperforms benchmarks, including asset-specific models, return correlation-based graphs, and graph structures based on historical PnL or Sharpe Ratio attained by single-asset generators.

Tails of Cross-Sectional Return Distributions at High Frequencies

T. G. Andersen, Y. Ding, V. Todorov
Northwestern University, Macau University

We develop nonparametric estimates for tail risk in the cross-section of asset prices at high frequencies. We show that the tail behavior of the cross-sectional return distribution depends on whether the time interval contains a systematic jump event. If so, the cross-sectional return tail is governed by the assets' exposures to the systematic event while, otherwise, it is determined by the idiosyncratic jump tails of the stocks. We develop an estimator for the tail shape of the cross-sectional return distribution that display distinct properties with and without systematic jumps. Empirically, we provide evidence for symmetric cross-sectional return tails at high-frequency that exhibit nontrivial and persistent time series variation. A hypothesis of equal cross-sectional return tail shapes during periods with and without systematic jump events is strongly rejected by the data.

SING: SDE Inference via Natural Gradients

A. Hu, H. Smith, S. Linderman
Stanford University

Latent stochastic differential equation (SDE) models are important tools for the unsupervised discovery of dynamical systems from data, with applications ranging from engineering to neuroscience. In these complex domains, exact posterior inference of the latent state path is typically intractable, motivating the use of approximate methods such as variational inference (VI). However, existing VI methods for inference in latent SDEs often suffer from slow convergence and numerical instability. Here, we propose SDE Inference via Natural Gradients (SING), a method that leverages natural gradient VI to efficiently exploit the underlying geometry of the model and variational posterior. SING enables fast and reliable inference in latent SDE models by approximating intractable integrals and parallelizing computations in time. We provide theoretical guarantees that SING will approximately optimize the intractable, continuous-time objective of interest. Moreover, we demonstrate that better state inference enables more accurate estimation of nonlinear drift functions using, for example, Gaussian process SDE models. SING outperforms prior methods in state inference and drift estimation on a variety of datasets, including a challenging application to modeling neural dynamics in freely behaving animals. Altogether, our results illustrate the potential of SING as a tool for accurate inference in complex dynamical systems, especially those characterized by limited prior knowledge and non-conjugate structure.

Efficient Difference-in-Differences and Event Study Estimators

X. Chen, P. H. C. Sant'Anna, H. Xie
Yale University, Emory University, Peking University

This paper investigates efficient Difference-in-Differences (DiD) and Event Study (ES) estimation using short panel data sets within the heterogeneous treatment effect framework, free from parametric functional form assumptions and allowing for variation in treatment timing. We provide an equivalent characterization of the DiD potential outcome model using sequential conditional moment restrictions on observables, which shows that the DiD identification assumptions typically imply nonparametric overidentification restrictions. We derive the semiparametric efficient influence function (EIF) in closed form for DiD and ES causal parameters under commonly imposed parallel trends assumptions. The EIF is automatically Neyman orthogonal and yields the smallest variance among all asymptotically normal, regular estimators of the DiD and ES parameters. Leveraging the EIF, we propose simple-to-compute efficient estimators. Our results highlight how to optimally explore different pre-treatment periods and comparison groups to obtain the tightest (asymptotic) confidence intervals, offering practical tools for improving inference in modern DiD and ES applications even in small samples. Calibrated simulations and an empirical application demonstrate substantial precision gains of our efficient estimators in finite samples.

Finance and Economics

Do Longer Block Times Impair Market Efficiency in Decentralized Markets?

A. Capponi, Á. Cartea, F. Drissi

COLUMBIA UNIVERSITY, UNIVERSITY OF OXFORD

In blockchains, traders bid priority fees to compete for queue position at regular intervals known as block time. Priority fees reduce trading volumes of informed traders and mitigate adverse selection costs. Longer block times can improve price efficiency by increasing speculative profits and by encouraging traders to acquire information. However, beyond a certain block time, expected losses from adverse selection outweigh the benefits of priority fees, leading to a liquidity freeze—unless uninformed liquidity demand grows sufficiently with block time. In equilibrium, traders only reveal their information immediately before the block is created.

Seemingly Virtuous Complexity in Return Prediction

S. NAGEL

UNIVERSITY OF CHICAGO, NBER, CEPR, AND CESIFO

Return prediction with Random Fourier Features (RFF)—a very large number, P, of nonlinear transformations of a small number, K, of predictor variables—has become popular recently. Surprisingly, this approach appears to yield successful out-of-sample stock market index return predictions even when trained in rolling windows as small as $T = 12$ months with P in the thousands, and without shrinkage. However, this apparent virtue of complexity is an illusion. When $P \gg T$, the RFF-based predictive regression closely approximates a kernel ridgeless regression that takes the original K predictors as input. The resulting forecast is effectively a weighted average of past returns, with weights highest on periods whose predictor vectors are most similar to the current one. In short training windows, similarity simply means recency, so the forecast reduces to a weighted average of recent returns—essentially a momentum strategy. Moreover, similarity of current and past predictor vectors is decreasing in predictor volatility, which makes it a volatility-timed momentum strategy. Crucially, the volatility-timed momentum nature of the strategy arises mechanically from predictor similarity. The strong historical performance of the RFF-based strategy thus reflects the coincidental historical success of volatility-timed momentum, not predictive information extracted from training data.

Understanding the Virtue of Complexity

B. T. KELLY, S. MALAMUD

YALE UNIVERSITY, AQR CAPITAL MANAGEMENT, NBER, SWISS FINANCE INSTITUTE, EPFL, AND CEPR

Recent papers have challenged certain aspects of the "virtue of complexity" described by Kelly et al. (2024) (KMZ) and related work. These challenges ultimately have little bearing on the theoretical arguments or empirical findings of KMZ. They do, however, provide a valuable opportunity to better understand the nuanced behavior of complex models. In addition to responding to recent challenges, we provide detailed discussions of how complex models learn in small samples, the roles of "nominal" and "effective" complexity, the unique effects of implicit regularization, and the importance of limits to learning. We then present new empirical and theoretical analyses that expand on KMZ. Finally, we introduce and demonstrate the virtue of ensemble complexity.

Anomalies as New Hedge Fund Factors

Y. CHEN, S. Z. LI, Y. TANG, G. ZHOU

TEXAS A&M UNIVERSITY, RUTGERS UNIVERSITY, SHANGHAI UNIVERSITY OF FINANCE AND ECONOMICS, WASHINGTON UNIVERSITY IN ST. LOUIS

We identify a parsimonious set of factors from a large set of candidates for explaining hedge fund returns, ranging from equity market, anomaly and trend-following factors to macroeconomic factors. The resulting nine-factor model, including five anomaly factors, outperforms existing hedge fund models both in-sample and out-of-sample, with a significant reduction in alphas while showing substantial cross-sectional performance heterogeneity. Further analysis based on fund holdings confirms the model's ability to capture returns from arbitrage trading. Overall, the anomaly factors help quantify hedge fund strategies and risk exposures and improve fund performance evaluation.

Bitcoin Price Volatility: Effects of Retail Traders, Illegal Users, and Sentiment

K. JOHN, J. LI

NEW YORK UNIVERSITY, STEVENS INSTITUTE OF TECHNOLOG

In this paper, we study the determinants of Bitcoin volatility. We estimate and decompose Bitcoin's realized volatility into jump volatility and continuous volatility components. We find that innovations in Robinhood retail trading are positively related to the continuous component of Bitcoin's volatility. Specifically, the innovations in Robinhood retail trading predict an increase in future ten-day continuous volatility. The innovations in the (anonymous) trading volume of Monero are also positively related to the jump component in Bitcoin's volatility. Specifically, the innovations in Monero trading volume predict an increase in average jump volatility over the next five days. Our results suggest that retail trading positively influences the continuous component of Bitcoin volatility, while the preference for transaction anonymity positively affects the jump component of Bitcoin volatility.

Factor Model Selection Using the ICAPM

H. MAMAYSKY, P. GLASSERMAN, J. HUBER

Columbia Business School

Using market data together with household-level consumption data, we extend the factor model ICAPM consistency test of Maio and Santa-Clara (2012). We find that more consistent factor models have less persistent alphas, and more stable betas and out-of-sample mean squared errors. We propose a novel statistical test for the sign of the consistency-stability relationship across many factor models and over time. Our methodology allows for the efficient identification of the historically most ICAPM-consistent factor models and factors from a large candidate pool. Our results suggest that historically consistent models are likely to be stable in the future.

Failing Banks

S. Correia, S. Luck, E. Verner

Federal Reserve Bank of Richmond, Federal Reserve Bank of New York, MIT Sloan School of Management and NBER

Why do banks fail? We create a panel covering most commercial banks from 1863 through 2024 to study the history of failing banks in the United States. Failing banks are characterized by rising asset losses, deteriorating solvency, and an increasing reliance on expensive noncore funding. These commonalities imply that bank failures are highly predictable using simple accounting metrics from publicly available financial statements. Failures with runs were common before deposit insurance, but these failures are strongly related to weak fundamentals, casting doubt on the importance of non-fundamental runs. Furthermore, low recovery rates on failed banks' assets suggest that most failed banks were fundamentally insolvent, barring strong assumptions about the value destruction of receiverships. Altogether, our evidence suggests that the primary cause of bank failures and banking crises is almost always and everywhere a deterioration of bank fundamentals.

Cursed Equilibria and Knightian Uncertainty in a Trading Game

J. Preker

Bielefeld University

We introduce a novel equilibrium concept that incorporates Knightian uncertainty into the cursed equilibrium (Eyster and Rabin, 2005). This concept is then applied to a two-player game in which agents can engage in trade or refuse to do so. While the Bayesian Nash equilibrium predicts that trade never happens, players do trade in a cursed equilibrium. The inclusion of uncertainty enhances this effect for cursed and uncertainty averse players. This contrasts general findings that uncertainty reduces trade but is consistent with behavior that has been observed in experiments.

Non-Discriminatory Personalized Pricing

P. Strack, K. H. Yang

Yale University

A monopolist offers personalized prices to consumers with unit demand, heterogeneous values, and idiosyncratic costs, who differ in a protected characteristic, such as race or gender. The seller is subject to a non-discrimination constraint: consumers with the same cost, but different characteristics must face identical prices. Such constraints arise in regulated markets like credit or insurance. The setting reduces to an optimal transport, and we characterize the optimal pricing rule. Under this rule, consumers may retain surplus, and either group may benefit. Strengthening the constraint to cover transaction prices redistributes surplus, harming the low-value group and benefiting the high-value group.

Mathematical and Computational Finance

Optimal Dynamic Fees in Automated Market Makers

L. Baggiani, M. Herdegen, L. Sánchez-Betancourt

University of Warwick, University of Stuttgart, University of Oxford

Automated Market Makers (AMMs) are emerging as a popular decentralised trading platform. In this work, we determine the optimal dynamic fees in a constant function market maker. We find approximate closed-form solutions to the control problem and study the optimal fee structure. We find that there are two distinct fee regimes: one in which the AMM imposes higher fees to deter arbitrageurs, and another where fees are lowered to increase volatility and attract noise traders. Our results also show that dynamic fees that are linear in inventory and are sensitive to changes in the external price are a good approximation of the optimal fee structure and thus constitute suitable candidates when designing fees for AMMs.

Local Signature-Based Expansions

F. M. Bandi, R. Renò, S. Svaluto-Ferro

Johns Hopkins University, ESSEC Business School, University of Verona

We study the local (in time) expansion of a continuous-time process and its conditional moments, including the process' characteristic function. The expansions are conducted by using the properties of the (time-extended) Ito signature, a tractable basis composed of iterated integrals of the driving deterministic and stochastic signals: time, multiple correlated Brownian motions and multiple correlated compound Poisson processes. We show that these properties are conducive to automated expansions to any order with explicit

coefficients and, therefore, to stochastic representations in which asymptotics can be conducted for a shrinking time (t to 0), as in the extant continuous-time econometrics literature, but, also, for a fixed time (such that t smaller than 1) with a diverging expansion order. The latter design opens up novel opportunities for identifying deep characteristics of the assumed process.

Designing Funding Rates for Perpetual Futures in Cryptocurrency Markets

J. Kim, H. Park
Seoul National University

In cryptocurrency markets, a key challenge for perpetual future issuers is maintaining alignment between the perpetual future price and target value. This study addresses this challenge by exploring the relationship between funding rates and perpetual future prices. Our results demonstrate that by appropriately designing funding rates, the perpetual future price can remain aligned with the target value. We develop replicating portfolios for perpetual futures, offering issuers an effective method to hedge their positions. Additionally, we provide path-dependent funding rates as a practical alternative and investigate the difference between the original and path-dependent funding rates. To achieve these results, our study employs path-dependent infinite-horizon BSDEs in conjunction with arbitrage pricing theory. Our main results are obtained by establishing the existence and uniqueness of solutions to these BSDEs and analyzing the large-time behavior of these solutions.

Split the Yield, Share the Risk: Pricing, Hedging and Fixed Rates in DeFi

V. Nadkarni, P. Viswanath
Princeton University

We present the first formal treatment of *yield tokenization*, a mechanism that decomposes yield-bearing assets into principal and yield components to facilitate risk transfer and price discovery in decentralized finance (DeFi). We propose a model that characterizes yield token dynamics using stochastic differential equations. We derive a no-arbitrage pricing framework for yield tokens, enabling their use in hedging future yield volatility and managing interest rate risk in decentralized lending pools. Taking DeFi lending as our focus, we show how both borrowers and lenders can use yield tokens to achieve optimal hedging outcomes and mitigate exposure to adversarial interest rate manipulation. Furthermore, we design automated market makers (AMMs) that incorporate a menu of bonding curves to aggregate liquidity from participants with heterogeneous risk preferences. This leads to an efficient and incentive-compatible mechanism for trading yield tokens and yield futures. Building on these foundations, we propose a modular *fixed-rate* lending protocol that synthesizes on-chain yield token markets and lending pools, enabling robust interest rate discovery and enhancing capital efficiency. Our work provides the theoretical underpinnings for risk management and fixed-income infrastructure in DeFi, offering practical mechanisms for stable and sustainable yield markets.

Acknowledgement

The OMI Research Newsletter is curated by Álvaro Cartea, Director of the Oxford-Man Institute, and Shifan Yu, postdoctoral researcher at the institute. Paper suggestions are from fellow members at the institute: Álvaro Arroyo, Patrick Chang, Fayçal Drissi, Gerardo Durán-Martín, and Qi Jin.

References

Andersen, Torben G, Yi Ding, and Viktor Todorov. 2025. "Tails of Cross-Sectional Return Distributions at High Frequencies." *SSRN Working Paper*.

Baggiani, Leonardo, Martin Herdegen, and Leandro Sánchez-Betancourt. 2025. "Optimal Dynamic Fees in Automated Market Makers." *arXiv preprint arXiv:2506.02869*.

Bandi, Federico M, Roberto Renò, and Sara Svaluto-Ferro. 2025. "Local Signature-Based Expansions." *arXiv preprint arXiv:2504.06351*.

Capponi, Agostino, Alvaro Cartea, and Fayçal Drissi. 2025. "Do Longer Block Times Impair Market Efficiency in Decentralized Markets?" *SSRN Working Paper*.

Chen, Xiaohong, Pedro HC Sant'Anna, and Haitian Xie. 2025. "Efficient Difference-in-Differences and Event Study Estimators." *arXiv preprint arXiv:2506.17729*.

Chen, Yong, Sophia Zhengzi Li, Yushan Tang, and Guofu Zhou. 2025. "Anomalies as New Hedge Fund Factors." *Journal of Financial and Quantitative Analysis*, forthcoming.

Correia, Sergio, Stephan Luck, and Emil Verner. 2025. "Failing Banks." *arXiv preprint arXiv:2506.06082*.

Duran-Martin, Gerardo, Leandro Sánchez-Betancourt, Álvaro Cartea, and Kevin Murphy. 2025. "Scalable Generalized Bayesian Online Neural Network Training for Sequential Decision Making." *arXiv preprint arXiv:2506.11898*.

Hu, Amber, Henry Smith, and Scott Linderman. 2025. "SING: SDE Inference via Natural Gradients." *arXiv preprint arXiv:2506.17796*.

John, Kose, and Jingrui Li. 2025. "Bitcoin Price Volatility: Effects of Retail Traders, Illegal Users, and Sentiment." *SSRN Working Paper*.

Kelly, Bryan T, and Semyon Malamud. 2025. "Understanding the Virtue of Complexity." *SSRN Working Paper*.

Kim, Jaehyun, and Hyungbin Park. 2025. "Designing Funding Rates for Perpetual Futures in Cryptocurrency Markets." *arXiv preprint arXiv:2506.08573*.

Mamaysky, Harry, Paul Glasserman, and Jeremias Huber. 2025. "Factor Model Selection Using the ICAPM." *SSRN Working Paper*.

Nadkarni, Viraj, and Pramod Viswanath. 2025. "Split the Yield, Share the Risk: Pricing, Hedging and Fixed Rates in DeFi." *arXiv preprint arXiv:2505.22784*.

Nagel, Stefan. 2025. "Seemingly Virtuous Complexity in Return Prediction." *SSRN Working Paper*.

Preker, Jurek. 2025. "Cursed Equilibria and Knightian Uncertainty in a Trading Game." *arXiv preprint arXiv:2506.10663*.

Strack, Philipp, and Kai Hao Yang. 2025. "Non-Discriminatory Personalized Pricing." *arXiv preprint arXiv:2506.20925*.

Vuletić, Milena, and Mihai Cucuringu. 2025. "GraFiN-Gen: Graph-Based Ensemble Generative Modelling for Multi-Asset Forecasting." *SSRN Working Paper*.