VOL. 5, NO. 3

OMI Research Newsletter – December 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 two full-time OMI-Hooke Postdoctoral Research Fellows in Quantitative Finance, jointly with the Mathematical Institute. Please find more details here.

Highlights

Key contributions in this month's edition: Wang, Richardson and Robins (2025) offer a systematic comparative analysis of three main frameworks for causal inference and show how they could be coherently integrated within empirical research. Goldsmith-Pinkham and Lyu (2025) recast financial event studies in a modern causal framework, showing that standard factor-model abnormal returns are generally inconsistent and that synthetic control methods more robustly identify average treatment effects. Chen et al. (2025) develop an "Econometrics AI Agent" that automates applied econometric workflows and shows markedly higher accuracy than general-purpose LLMs and agents. Cartea et al. (2025) show that realistic stock-level liquidity constraints severely limit the scalable profits and Sharpe ratios of anomaly-based trading strategies. Cartea, Cucuringu and Kim (2025) train a self-supervised tabular transformer to learn firm-characteristic embeddings and then builds similarity-based long–short portfolios that outperform standard factor and machine-learning models in Sharpe ratio and return prediction. O'Doherty, Wang and Yan (2025) show that many macroeconomic variables—especially NIPA and housing factors—earn significant risk premia and can explain most stock return anomalies, revealing strong macroeconomic foundations for the "anomaly zoo." Chang et al. (2025) find that a few functional principal components of the S&P 500 risk-neutral density significantly forecast monthly market returns and enable profitable market-timing strategies. Cohen and Svosve (2025) use Assumed Density Filtering to turn Heston-style stochastic volatility models into lightweight path-dependent volatility models that fit both realized volatility and SPX/VIX smiles. Baggiani, Herdegen and Khan (2025) show that a one-period portfolio choice problem with a risk constraint has optimal solutions across all markets when either investors' utility or the risk measure is sensitive to large losses.

Statistics, Econometrics & Machine Learning

Causal Inference: A Tale of Three Frameworks

L. WANG, T. RICHARDSON, J. ROBINS
UNIVERSITY OF TORONTO, UNIVERSITY OF WASHINGTON, HARVARD UNIVERSITY

Causal inference is a central goal across many scientific disciplines. Over the past several decades, three major frameworks have emerged to formalize causal questions and guide their analysis: the potential outcomes framework, structural equation models, and directed acyclic graphs. Although these frameworks differ in language, assumptions, and philosophical orientation, they often lead to compatible or complementary insights. This paper provides a comparative introduction to the three frameworks, clarifying their connections, highlighting their distinct strengths and limitations, and illustrating how they can be used together in practice. The discussion is aimed at researchers and graduate students with some background in statistics or causal inference who are seeking a conceptual foundation for applying causal methods across a range of substantive domains.

Causal Inference in Financial Event Studies

P. Goldsmith-Pinkham, T. Lyu

Yale University, NBER

Financial event studies, ubiquitous in finance research, typically use linear factor models with known factors to estimate abnormal returns and identify causal effects of information events. This paper demonstrates that when factor models are misspecified – an almost certain reality – traditional event study estimators produce inconsistent estimates of treatment effects. The bias is particularly severe during volatile periods, over long horizons, and when event timing correlates with market conditions. We derive precise conditions for identification and expressions for asymptotic bias. As an alternative, we propose synthetic control methods that construct replicating portfolios from control securities without imposing specific factor structures. Revisiting four empirical applications, we show that some established findings may reflect model misspecification rather than true treatment effects. While traditional methods remain reliable for short-horizon studies with random event timing, our results suggest caution when interpreting long-horizon or volatile-period event studies and highlight the importance of quasi-experimental designs when available.

Can AI Master Econometrics? Evidence from Econometrics AI Agent on Expert-Level Tasks

Q. Chen, T. Han, J. Li, Y. Luo, Y. Wu, X. Zhang, T. Zhou

Shandong University, The University of Hong Kong, Hong Kong University of Science and Technology

Can AI effectively perform complex econometric analysis traditionally requiring human expertise? This paper evaluates AI agents' capability to master econometrics, focusing on empirical analysis performance. We develop an "Econometrics AI Agent" built on the open-source MetaGPT framework. This agent exhibits outstanding performance in: (1) planning econometric tasks strategically, (2) generating and executing code, (3) employing error-based reflection for improved robustness, and (4) allowing iterative refinement through multi-round conversations. We construct two datasets from academic coursework materials and published research papers to evaluate performance against real-world challenges. Comparative testing shows our domain-specialized AI agent significantly outperforms both benchmark large language models (LLMs) and general-purpose AI agents. This work establishes a testbed for exploring AI's impact on social science research and enables cost-effective integration of domain expertise, making advanced econometric methods accessible to users with minimal coding skills. Furthermore, our AI agent enhances research reproducibility and offers promising pedagogical applications for econometrics teaching.

Robust Reinforcement Learning under Diffusion Models for Data with Jumps

C. Jiang, D. Kim, A. Quintos, Y. Wang

University of Wisconsin-Madison, University of California, Riverside

Reinforcement Learning (RL) has proven effective in solving complex decision-making tasks across various domains, but challenges remain in continuous-time settings, particularly when state dynamics are governed by stochastic differential equations (SDEs) with jump components. In this paper, we address this challenge by introducing the Mean-Square Bipower Variation Error (MSBVE) algorithm, which enhances robustness and convergence in scenarios involving significant stochastic noise and jumps. We first revisit the Mean-Square TD Error (MSTDE) algorithm, commonly used in continuous-time RL, and highlight its limitations in handling jumps in state dynamics. The proposed MSBVE algorithm minimizes the mean-square quadratic variation error, offering improved performance over MSTDE in environments characterized by SDEs with jumps. Simulations and formal proofs demonstrate that the MSBVE algorithm reliably estimates the value function in complex settings, surpassing MSTDE's performance when faced with jump processes. These findings underscore the importance of alternative error metrics to improve the resilience and effectiveness of RL algorithms in continuous-time frameworks.

Do We Need Dozens of Methods for Real World Missing Value Imputation?

K. Grzesiak, C. Muller, J. Josse, J. Näf

University of Wrocław, University of Oxford, University of Montpellier, University of Geneva

Missing values pose a persistent challenge in modern data science. Consequently, there is an ever-growing number of publications introducing new imputation methods in various fields. While many studies compare imputation approaches, they often focus on a limited subset of algorithms and evaluate performance primarily through pointwise metrics such as RMSE, which are not suitable to measure the preservation of the true data distribution. In this work, we provide a systematic benchmarking method based on the idea of treating imputation as a distributional prediction task. We consider a large number of algorithms and, for the first time, evaluate them not only on synthetic missing mechanisms, but also on real-world missingness scenarios, using the concept of Imputation Scores. Finally, while the focus of previous benchmark has often been on numerical data, we also consider mixed data sets in our study. The analysis overwhelmingly confirms the superiority of iterative imputation algorithms, especially

the methods implemented in the mice R package.

Finance and Economics

Bottom-Up Capacity Constraints and the Limits of Anomaly Profitability

Á. Cartea, M. Cucuringu, Q. Jin, J. Zhu

Oxford-Man Institute for Quantitative Finance, University of Oxford, University of California, Los Angeles

We propose a new bottom-up framework to evaluate the implementable value of trading strategies. Our approach is the first to assess profitability in terms of dollar trade sizes and dollar returns, thereby assessing the feasible trading capacity. Applying the framework to a broad set of asset pricing anomalies and return prediction models, we find that the realized Sharpe ratio of anomaly-based strategies declines substantially when liquidity constraints are considered. Although some anomalies and models appear highly predictive in statistical terms, their predictive power is concentrated in illiquid stocks where tradable volume is limited. While returns are further reduced when one accounts for trading costs, liquidity constraints are the primary trading friction that largely limits the trading capacity of the anomaly-based strategies. Our findings reveal a divergence between predictability and profitability and highlight the need to evaluate trading strategies by their implementable value and dollar returns under realistic liquidity conditions.

Learning Firm Characteristics for Asset Management

Á. Cartea, M. Cucuringu, D. Kim

Oxford-Man Institute for Quantitative Finance, University of Oxford, University of California, Los Angeles

We propose a novel approach for constructing portfolios with learned representations of firm characteristics. Our framework introduces a multi-task self-supervised learning framework to train a tabular transformer with historical data (e.g., changes in earnings and momentum for a stock) to capture cross-sectional and temporal patterns in firm fundamentals. The model summarizes the data into lower-dimensional embeddings which we use to build similarity-based investment strategies. The Sharpe ratio, cumulative returns, and the accuracy of predicted returns of our strategy outperform standard benchmarks. Finally, we examine the key firm characteristics that the transformer identifies as most influential over time and during recessions.

On the Macroeconomic Foundations of the Anomaly Zoo

M. S. O'Doherty, F. Wang, S. Yan

University of Missouri, Miami University, Lehigh University

We apply modern asset pricing methods that mitigate omitted variable and measurement error biases to estimate risk premia for 190 candidate macroeconomic factors using a broad cross section of equity style portfolios. More than 40 macroeconomic factors carry statistically significant risk premia. Models that include tradable mimicking portfolios for these factors frequently outperform leading multifactor models in explaining CAPM anomalies. Our findings reveal a strong link between economic fluctuations and asset prices, with the empirically most impressive factors tied to NIPA aggregates and housing market activity.

Market Returns Dormant in Option Panels

Y. Chang, Y. Choi, S. Kim, J. Park

Indiana University, Xavier University, Korea Advanced Institute of Science and Technology

This paper develops a novel functional predictive regression framework linking option-implied distributions to stock market returns, motivated by the fundamental link between risk-neutral and physical densities. Using SP 500 option panels, our model exhibits significant forecasting power, achieving robust out-of-sample R^2 exceeding 4% for monthly return predictions, outperforming traditional predictors. Superior performance arises from leveraging the full spectrum of the risk-neutral density via functional principal components. Our analysis reveals that forecasting success stems from nuanced variations in risk-neutral densities, reflecting various economic states of risk perception and trading frictions, and demonstrates potential economic gains through a markettiming strategy.

Model-Free Factor Risk Premia

M. Kazemi

Arizona State University

I develop a semi-parametric estimator for factor risk premia that is robust to model misspecification. Standard two-pass and GMM estimators produce biased estimates when priced factors are omitted, irrelevant factors are included, or factors are weakly identified. The proposed method avoids these problems by separating the construction of the stochastic discount factor from any factor model specification. The SDF is recovered by minimizing Kullback-Leibler divergence subject only to no-arbitrage restrictions, and each factor's premium is computed through the marginal moment–$R_f E[mf]$. This approach ensures that the estimate for any given factor depends only on that factor and the recovered SDF-not on which other factors the researcher specifies. Monte Carlo

simulations calibrated to Fama-French-Carhart factors show that the estimator matches conventional methods under correct specification and delivers root mean squared errors 90-99% smaller when priced factors are omitted. Empirically, exponential-tilting estimates for traded factors align closely with timeseries means, while Fama-MacBeth estimates for non-traded macro factors exhibit wild specification sensitivity-with sign reversals and swings exceeding 30 percentage points-that disappear under the robust approach.

Innovation-Driven Contractions: A Missing Link for Asset Pricing Puzzles

G. Segal, C. Ying

University of North Carolina at Chapel Hill, The Chinese University of Hong Kong

We provide a unified resolution for prominent conditional asset-pricing puzzles through "innovation-driven contractions." Empirically, we document two stylized facts: (1) technological innovations initially cause significant labor contractions and have negligible impact on capital, and (2) puzzling disconnections-between stock returns and either investment growth or hiring surprises-intensify under stickier prices. A two-sector New-Keynesian model quantitatively ties these facts, explaining the anomalous negative correlation between stock-returns and investment-returns, the negative market reactions to favorable labor market news, and the procyclical equity yield term-structure slope. Investment-based dividend yields filtered from our model strongly predict future returns, underscoring our mechanism's empirical applicability.

The Economic Impact of Brexit

N. Bloom, P. Bunn, P. Mizen, P. Smietanka, G. Thwaites

Stanford University, NBER, Deutsche Bundesbank, King's College London, Bank of England, University of Nottingham

This paper examines the impact of the UK's decision to leave the European Union (Brexit) in 2016. Using almost a decade of data since the referendum, we combine simulations based on macro data with estimates derived from micro data collected through our Decision Maker Panel survey. These estimates suggest that by 2025, Brexit had reduced UK GDP by 6% to 8%, with the impact accumulating gradually over time. We estimate that investment was reduced by between 12% and 18%, employment by 3% to 4% and productivity by 3% to 4%. These large negative impacts reflect a combination of elevated uncertainty, reduced demand, diverted management time, and increased misallocation of resources from a protracted Brexit process. Comparing these with contemporary forecasts – providing a rare macro example to complement the burgeoning micro-literature of social science predictions – shows that these forecasts were accurate over a 5-year horizon, but they underestimated the impact over a decade.

Mathematical and Computational Finance

Linking Path-Dependent and Stochastic Volatility Models

S. N. Cohen, C. Svosve

University of Oxford, Oxford-Man Institute for Quantitative Finance

We explore a link between stochastic volatility (SV) and path-dependent volatility (PDV) models. Using assumed density filtering, we map a given SV model into a corresponding PDV representation. The resulting specification is lightweight, improves in-sample fit, and delivers robust out-of-sample forecasts. We also introduce a calibration procedure for both SV and PDV models that produces standard errors for parameter estimates and supports joint calibration of SPX/VIX smile.

The Interplay between Utility and Risk in Portfolio Selection

L. Baggiani, M. Herdegen, N. Khan

University of Warwick, University of Oxford, Oxford-Man Institute for Quantitative Finance

We revisit the problem of portfolio selection, where an investor maximizes utility subject to a risk constraint. Our framework is very general and accommodates a wide range of utility and risk functionals, including non-concave utilities such as S-shaped utilities from prospect theory and non-convex risk measures such as Value at Risk. Our main contribution is a novel and complete characterization of well-posedness for utility-risk portfolio selection in one period that takes the interplay between the utility and the risk objectives fully into account. We show that under mild regularity conditions the minimal necessary and sufficient condition for well-posedness is given by a very simple eitheror criterion: either the utility functional or the risk functional need to satisfy the axiom of sensitivity to large losses. This allows to easily describe well-posedness or ill-posedness for many utility/risk pairs, which we illustrate by a large number of examples. In the special case of expected utility maximization without a risk constraint (but including non-concave utilities), we show that well-posedness is fully characterised by the asymptotic loss-gain ratio, a simple and interpretable quantity that describes the investor's asymptotic relative weighting of large losses versus large gains.

CoFinDiff: Controllable Financial Diffusion

Model for Time Series Generation

Y. Tanaka, R. Hashimoto, T. Takayanagi, Z. Piao, Y. Murayama, K. Izumi

The University of Tokyo, Nomura Holdings, Inc.

The generation of synthetic financial data is a critical technology in the financial domain, addressing challenges posed by limited data availability. Traditionally, statistical models have been employed to generate synthetic data. However, these models fail to capture the stylized facts commonly observed in financial data, limiting their practical applicability. Recently, machine learning models have been introduced to address the limitations of statistical models; however, controlling synthetic data generation remains challenging. We propose CoFinDiff (Controllable Financial Diffusion model), a synthetic financial data generation model based on conditional diffusion models that accept conditions about the synthetic time series. By incorporating conditions derived from price data into the conditional diffusion model via cross-attention, CoFinDiff learns the relationships between the conditions and the data, generating synthetic data that align with arbitrary conditions. Experimental results demonstrate that: (i) synthetic data generated by CoFinDiff capture stylized facts; (ii) the generated data accurately meet specified conditions for trends and volatility; (iii) the diversity of the generated data surpasses that of the baseline models; and (iv) models trained on CoFinDiff-generated data achieve improved performance in deep hedging task.

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: Patrick Chang, Fayçal Drissi, Gerardo Duran-Martin, Christopher Hyland, Jiawen Liang, and Kevin Schneider.

References

Baggiani, Leonardo, Martin Herdegen, and Nazem Khan. 2025. "The Interplay between Utility and Risk in Portfolio Selection." *arXiv preprint arXiv:2509.10351*.

Bloom, Nicholas, Philip Bunn, Paul Mizen, Pawel Smietanka, and Gregory Thwaites. 2025. "The Economic Impact of Brexit." *NBER Working Paper 34459*.

Cartea, Álvaro, Mihai Cucuringu, and Donggeun Kim. 2025. "Learning Firm Characteristics for Asset Management." *Available at SSRN 5853103*.

Cartea, Álvaro, Mihai Cucuringu, Qi Jin, and Jiexiu Zhu. 2025. "Bottom-Up Capacity Constraints and the Limits of Anomaly Profitability." *Available at SSRN 5797502*.

Chang, Yoosoon, Youngmin Choi, Soohun Kim, and Joon Park. 2025. "Market Returns Dormant in Option Panels." *Available at SSRN 5632350*.

Chen, Qiang, Tianyang Han, Jin Li, Ye Luo, Yuxiao Wu, Xiaowei Zhang, and Tuo Zhou. 2025. "Can AI Master Econometrics? Evidence from Econometrics AI Agent on Expert-Level Tasks." *arXiv preprint arXiv:2506.00856*.

Cohen, Samuel N, and Cephas Svosve. 2025. "Linking Path-Dependent and Stochastic Volatility Models." *arXiv preprint arXiv:2510.02024*.

Goldsmith-Pinkham, Paul, and Tianshu Lyu. 2025. "Causal Inference in Financial Event Studies." *arXiv preprint arXiv:2511.15123*.

Grzesiak, Krystyna, Christophe Muller, Julie Josse, and Jeffrey Näf. 2025. "Do we Need Dozens of Methods for Real World Missing Value Imputation?" *arXiv preprint arXiv:2511.04833*.

Jiang, Chenyang, Donggyu Kim, Alejandra Quintos, and Yazhen Wang. 2024. "Robust Reinforcement Learning under Diffusion Models for Data with Jumps." *arXiv preprint arXiv:2411.11697*.

Karagiorgis, Ariston, Dimitrios Anastasiou, Konstantinos Drakos, and Steven Ongena. 2025. "The Leverage of Hedge Funds and the Risk of Their Prime Brokers." *Available at SSRN 5312471*.

Kazemi, Maziar. 2025. "Model-Free Factor Risk Premia." *Available at SSRN 5686704*.

O'Doherty, Michael S, Feifei Wang, and Sterling Yan. 2025. "On the Macroeconomic Foundations of the Anomaly Zoo." *Available at SSRN 5395852*.

Segal, Gill, and Chao Ying. 2025. "Innovation-Driven Contractions: A Missing Link for Asset Pricing Puzzles." *Available at SSRN 5688729*.

Tanaka, Yuki, Ryuji Hashimoto, Takehiro Takayanagi, Zhe Piao, Yuri Murayama, and Kiyoshi Izumi. 2025. "CoFinDiff: Controllable Financial Diffusion Model for Time Series Generation." *arXiv preprint arXiv:2503.04164*.

Wang, Linbo, Thomas Richardson, and James Robins. 2025. "Causal Inference: A Tale of Three Frameworks." *arXiv preprint arXiv:2511.21516*.