Summary of *AI-POWERED (FINANCE) SCHOLARSHIP*

Robert Novy-Marx Mihail Z. Velikov (NBER, 2025)

Summarized by Li Ziming

1. **What are the research questions?**
   - How LLMs used to automatically generating academic finance papers?
   - What kind of threat will it pose to traditional academic norms?
2. **Why are the research questions interesting?**
   - Inherent tension between idealized scientific method and practical discovery. Hypothesizing often after results are known, but academic standard insists first develop theories and then test against data.
   - Emergence of LLMs transform the contradiction to technological capability. LLMs improve scientific reasoning and prediction, automatically generate research ideas, and conduct experiments and test hypotheses.
3. **What is the paper's contribution?**
   - Contribute to literature comparing AI-generated results and economic theories.
     - Prior literature: data mining has predictability comparable traditional peer review (Chen, Lopez-Lira, and Zimmermann, 2024).
     - Extend: LLM can rapidly produce coherent theoretical explanations for mined results.
   - Contribute to literature on application of LLM in hypotheses generation.
     - Prior literature: reason and predict (van Inwegen et al., 2023); generate hypotheses (Manning et al., 2024); conduct experiments and produce papers (Lu et al., 2024).
     - Extend: demonstrate a complete pipeline for automated academic research production in finance, from hypothesis generation through full paper creation.
4. **Sample: comment on the appropriateness of the sample selection procedures.**
   - Initial dataset contains 31,460 accounting variables and their temporal difference. Suitable for simple data mining.
5. **Comment on the appropriateness of variable definition and measurement.**
   - Signal names and acronyms are descriptive and creative.
   - Generated introduction, data section and conclusion similar to academic papers.
6. **Comment on the appropriateness of the regress/predict model specification.**
   - Implement a series of data-quality and sufficiency filters and increasingly stringent statistical validation tests to find statistically significant pattern.
7. **What difficulties arise in drawing inferences from the empirical work?**
   - Citation accuracy: fictitious references when cite specific or recent work.
   - Expression alignment: hypotheses, empirical results, description in each section.
8. **Describe at least one publishable and feasible extension of this research.**
   - Can it pass popular AIGC test?
   - How to design effective instructions for LLMs?

9. **What is the association between three literature?**

   - The first paper studies the use of large language models to automatically generate full academic papers.
   - The second paper proposes that large language models can use structural causal models to automatically generate hypotheses, design experiments, and run experimental simulations for different scenarios.
   - The third paper focuses on marketing experiments, using large language models to generate a large number of hypotheses, using machine learning algorithms to screen hypotheses, and then conducting experiments to verify the hypotheses.