



THE UNIVERSITY OF
CHICAGO

DEPARTMENT OF STATISTICS

Master's Thesis Presentation

Yilun Cai

Department of Statistics
The University of Chicago

“Empirical Asset Pricing via Transformer”

May 22, 2026, at 10:30 AM
Jones 111, 5747 S. Ellis Avenue

Abstract

This report describes a full empirical asset-pricing pipeline that combines the stock-characteristic data design of Gu et al. (2020), Empirical Asset Pricing via Machine Learning, with the stochastic-discount-factor and attention-based modeling framework of Kelly et al. (2025), Artificial Intelligence Asset Pricing Models. The report examines four model families: (i) a stock-level ridge return-prediction model, (ii) a linear stochastic discount factor (SDF), (iii) a linear transformer SDF that introduces cross-asset information sharing through linear attention, and (iv) a nonlinear deep transformer SDF and summarizes their performances as well as interpretability.