

NATHAN ZHANG

Vancouver, BC | nathanzhang1017@gmail.com | LinkedIn | GitHub | Website

RESEARCH EXPERIENCE

Research Assistant, UBC Sauder School of Business (Apr 2026 – Present)

Supervisors: J. Frank Li and Cheng Chen

- As sole RA on a study of age-biased technological change, link CPS, O*NET, and USPTO data, harmonizing 60+ historical O*NET releases. Construct PatentsBERTa exposure measures, validate them with Qwen 3 and sampled manual reading, and fit weighted FE OLS/GMM wage and sorting models.
- Run large-scale estimation, bootstrap inference, calibration, and robustness checks through R and Python workflows on HPC/Slurm. Maintain the replication package and draft the empirical section.

Research Assistant, UBC Sauder School of Business (May 2026 – Aug 2026)

Supervisor: Allen Hu

- Assessed RavenPack financial-news coverage against a ProQuest corpus of over 3.2 million articles, supporting research on financial information across media.
- Independently built reproducible yt-dlp ETL for multimodal financial research, generating panels from metadata, archiving and deduplicating video and audio, and validating subtitle availability.

Research Assistant, Vancouver School of Economics at UBC (Jan 2025 – May 2026)

Supervisors: Jonathan Graves and Laura Nelson

- Developed and deployed notebooks and HTML courseware for economics and social science students through COMET and Praxis UBC, applying computer vision, NLI, and sentiment analysis to historical and economic research. Led proposals and built image-embedding and classification workflows.
- Designed and taught a computer vision workshop. Co-developed a zero-shot NLI tutorial with human-label validation and robustness checks, in first revision at Programming Historian.

PUBLICATIONS AND MANUSCRIPTS

Memory Response Paths and Path-Structured Context Compilation (2026)

Joint first author with Jerrold Huang. Submitted to AAAI and under review

Proposed response path estimation to measure a memory record's incremental contribution across four horizons beyond endpoint scores. Designed paired-availability tests across 180 environments, fixed-evidence context ablations, model API workflows, and statistical analysis. In a separate 144-item MemGym-DR evaluation, PSSC improved LLM-judged quality over A-Mem, a prior SOTA conversational-memory framework, by 6.94 percentage points using 32.8% fewer API tokens.

Sequential Tone Residualization and Abnormal Returns in Earnings Calls (2026)

Joint first author with Jerrold Huang. Accepted at ICAIF 2026, selected for oral presentation

Developed context-adjusted earnings-call tone measures for investment signal research, co-designing information sets and building a time-aligned Capital IQ, IBES, and CRSP panel. Proposed the application and dictionary and LLM validation. Across 1,901 held-out 2025 calls, a one-SD answer-tone residual was associated with +61.5 bps in one-day market-adjusted return. Tone models achieved positive OOS R^2 , and prespecified joint pre-event placebo tests were nonsignificant.

EDUCATION

The University of British Columbia (Sep 2023 – May 2028)

B.A., Combined Major in Economics and Statistics, Minor in Data Science, Vancouver, BC

- Cumulative average: 93.5/100

RESEARCH INTERESTS

Methods: Statistical learning and inference, causal inference and econometrics, panel and time-series analysis, text and image methods.

Applications: Technological change and labor, financial risk and ESG, computational social science.

SELECTED PROJECTS

Double Machine Learning for Payment-Friction Spending Effects (2025)

Course research project [Code] [Report]

- Contributed EDA, leakage checks, multi-model refitting, and econometric input on estimands to a collaborative DoubleML analysis of payment friction and spending across about 5,600 respondents and roughly 22,000 in-person person-days. The study also explored heterogeneous effects with DR Learner.

Canadian Inflation Forecasting and Macro Signal Evaluation (2025)

Forecasting project [Code]

- Built an R and Quarto VAR forecast for Canadian inflation with rolling out-of-sample tests against naive and random-walk baselines. Presented impulse responses in a Bank of Canada-style policy briefing.

TEACHING AND RESEARCH INFRASTRUCTURE

Image Analysis Workshop: Computer Vision for Social Science Research (Mar 2026)

Workshop lead and instructor [Workshop]

Designed and taught image embeddings, classification, and detection for social science research.

A Study of Richter's Kouroi Through Image Embedding (Jul 2025)

Coauthor with Yash Mali and Krishaant Pathman [Notebook]

Used ConvNeXt V2 embeddings, PCA, clustering, and classification to test whether image embeddings recover structure in Richter's kouroi categories.

LLM Prediction Module (2025)

Coauthor [Module]

Coauthored course material on next-word prediction, token probabilities, and economics applications.

PROFESSIONAL EXPERIENCE

ESG Research Analyst Co-op, Capital Markets & Credit Investments (Sep 2026 – Present)

British Columbia Investment Management Corporation (BCI), Vancouver, BC

- Apply causal ML to study governance risk factors and financial performance for fundamental investment research, and develop ML models for M&A and other corporate events.
- Independently build reproducible Capital IQ and ISS ECA ETL in Databricks for voting-strategy backtests with simulated ballot-time information and subsequent-risk analysis.
- Assess portfolio environmental exposure using Net Purpose data.

Teaching Assistant (Sep 2024 – Aug 2025)

UBC Chinese Language Program, Vancouver, BC

- Supported Chinese language teaching, grading, volunteer training, and feedback coordination.

AWARDS AND SCHOLARSHIPS

- Work Learn International Undergraduate Research Awards (2026),
Economics Undergraduate Scholarship (2025), Faculty of Arts International Student Scholarship (2024, 2025)

SELECTED COURSEWORK

Statistical Inference (100), Machine Learning and Data Mining (93), Causal Methods (98)

Advanced Econometrics (98), Statistical Learning (97), Time Series and Forecasting (95)

TECHNICAL SKILLS AND LANGUAGES

Programming & data: Python, R, SQL, Bash, Julia, PySpark, DuckDB, Polars, BQL

Methods: econometrics, factor models, causal inference, panel and time-series, machine learning, NLP, computer vision

Modeling libraries: NumPy, SciPy, scikit-learn, PyTorch, statsmodels, fixest, OpenCV, spaCy, NLTK

Platforms & tools: Databricks, Bloomberg, DevOps and Git, GitHub Actions, HPC/Slurm, Jupyter and Quarto

Agent-assisted development: Codex, Claude Code

Languages: Mandarin Chinese: native, English: academic working proficiency, Japanese: JLPT N3