Boyuan Zhang

zhang.boyuan@hotmail.com | boyuan-zhang.github.io | boyuan-zhang415

Research Fields

Bayesian Econometrics, Time Series Analysis, Applied Macroeconomics, Machine Learning, Causal Inference, Foundation Models

Current Position

Amazon.com, Inc. Bellevue, WA
Economist II *2024 - Present*

Previous Positions

Amazon.com, Inc. Bellevue, WA
Postdoctoral Scientist *2023 - 2024*

Education

University of Pennsylvania 2017 - 2023
Ph.D. in Economics *GPA: 3.91/4*

University of Illinois at Urbana-Champaign 2014 - 2016
M.S. in Statistics *GPA: 3.97/4*

Renmin University of China 2010 - 2014
B.S. in Physics *GPA: 3.76/4*

Professional Experience

Amazon.com, Inc.

Economist II *Aug 2024 - Present*

- Developed and deployed Bayesian dynamic factor models for multivariate retail inventory forecasting across US, UK, EU, JP, and IN marketplaces, supporting quarterly financial planning and network design decisions.
  - Extended Canada-based framework to US market with multivariate inventory forecasting system, substantially outperforming incumbent model with more than 20% reduction in long-term forecast error.
  - Created explainability bridges (plan-vs-actuals, plan-over-plan, year-over-year) to decompose forecast changes and attribute them to specific input drivers, replacing black-box approaches.
  - Presented at CSS 2024 (Economics track, 21% acceptance rate) and delivered tech talk at Economist Summit.
- Quantified customer experience impacts and optimized fulfillment network planning using empirical causal inference methods to inform billions of dollars in investment decisions.
  - Built Bayesian nonparametric panel data model to measure heterogeneous delivery speed elasticities across US regions, identifying distinct groups with varying responses to local versus network-wide speed improvements (Presented at CSS 2024 (Lightning Talk) and AMLC 2024 (Poster, 30% acceptance)).
  - Evaluated same-day delivery network expansion using propensity score matching and difference-in-differences methods, quantifying significant incremental unit growth within weeks of new facility launches with heterogeneous effects by market maturity and baseline speed.
  - Developed selection optimization framework decomposing inventory strategy into breadth, assortment depth, and stock depth; demonstrated meaningful efficiency gains through counterfactual reallocation from low-performing product categories to high-demand items.
- Designed forward-looking customer segmentation and business analytics frameworks to drive strategic planning and operational improvements.
  - Developed BVAR-based dynamic customer spending segmentation incorporating external economic indicators (inflation, income, retail trends), enabling forward-looking bracket assignments and capturing **36%** more customers than existing ranking approaches with superior forecast accuracy. (Presented at CSS 2025 (Poster, 28%

acceptance) and Economist Summit 2025 (Poster, 27% acceptance)).
  - Analyzed FBA inventory positioning defects using dynamic panel regressions, identifying **$1.4B** revenue loss and quantifying root causes across seller performance and system operations.
  - Optimized low-velocity product selection using forward-looking demand forecasts, demonstrating over **$700M** in annualized free cash flow improvement potential from strategic inventory rationalization.
- Built operational analytics tools for supply chain visibility and risk mitigation across vendor lead time prediction and inventory imbalance monitoring.
  - Developed state-space model for vendor lead time backlog decomposition separating business-as-usual from defect-driven components with automated disruption detection, enabling proactive intervention and counterfactual performance benchmarking.
  - Measured regional inventory imbalance using dispersion metrics at product/region/weekly granularity, revealing negative correlation between total inventory and regional mismatch to inform network regionalization strategies.
- Served as Senior Area Chair for Forecasting, Macroeconomics & Finance track at Economics Summit 2025; reviewed submissions and selected papers for oral and poster presentations.

Postdoctoral Scientist *Aug 2023 - Aug 2024*

- Customized dynamic factor models and extracted factors from a wide panel of supply chain variables for accurate weekly inventory prediction in the US and Canada market, incorporating long-term trends, short-term deviation, seasonality, and holiday impacts.
- Applied a Bayesian framework for enhanced model inference, delivering meaningful predictive quantiles to inform fulfillment network design.
- Enhanced forecasting accuracy and efficiency by streamlining multi-stage processes, achieving substantial improvement in prediction density and quantiles.
- Incorporated business-driven prior assumptions for forecasting long-term behaviors and conducting scenario analyses under uncertainty.
- Implemented conditional forecasting to link actuals with predictions, improving error analysis and decision-making processes.
- Crafted an internal Python toolbox for implementing factor models, tailored for versatile applications in supply chain-related forecasting.

Economist Intern *May 2021 - Aug 2021*

- Developed and implemented structural time series models for the supply chain sector, producing precise weekly forecasts of key indicators in an area previously lacking specific weekly models.
- Achieved substantial improvement in out-of-sample prediction accuracy through comprehensive validation on both simulated and real-world datasets.
- Engineered an automated system for estimating the impact of floating holidays, and for identifying outliers, structural shifts, or one-time transient shocks in weekly data series.
- Conducted a comprehensive evaluation of various modeling approaches across econometrics, statistics, and machine learning, detailing their advantages and limitations.

SKILLS

Computing Skills

- **Languages**: Python, R, Julia, Matlab; SQL, STATA, SAS
- **Tools & Platforms**: AWS (S3, EC2, SageMaker), DataGrip, Git, LaTeX, Jupyter Notebooks
- **GenAI Tools**: Kiro, Kiro-CLI, Claude Code, Amazon Q

Modeling Skills

- **Econometrics**: structural vector autoregressions (SVAR), dynamic factor models, panel data models, Bayesian state-space models, quantile regression, nonparametric (kernel) regression, factor models, structural time series models, GMM, instrumental variables

- **Time Series Forecasting**: SARIMA, GARCH, exponential smoothing (ETS/Holt-Winters), STL decomposition, TBATS, Prophet, DeepAR, theta method, hierarchical reconciliation, ensemble forecasting methods
- **Machine Learning**: regularized regressions (LASSO, Ridge, Elastic Net), random forest, XGBoost, LightGBM, SVMs, neural networks (LSTM, Transformer-based), KMeans, KNN, PCA
- **Causal Inference**: difference-in-differences (DiD), staggered DiD, synthetic DiD, propensity score matching, regression discontinuity, synthetic control, double machine learning
- **Time Series Foundation Models**: Chronos, Chronos-2, TimesFM 2.5, Time-MoE, Moirai-2, Lag-Llama

PUBLICATIONS

"On the Aggregation of Probability Assessments: Regularized Mixtures of Predictive Densities for Eurozone Inflation and Real Interest Rates," *Journal of Econometrics*, 2023.

- Proposed methods for constructing regularized combination of density forecasts.
- Explored a variety of objectives and regularization penalties.
- Produced Eurozone inflation and real interest rate density forecasts.
- Showed better predictive performance than all individual inflation forecasters.

"When Will Arctic Sea Ice Disappear? Projections of Area, Extent, Thickness, and Volume," *Journal of Econometrics*, 2023.

- Provide point, interval, and density forecasts for four measures of Arctic Sea ice: area, extent, thickness, and volume.
- Enforced the joint constraint that the measures simultaneously arrive at an ice-free Arctic.
- Applied the constrained joint forecast procedure to models relating sea ice to atmospheric CO_2 concentration and models relating sea ice directly to time.
- Predicted an effectively ice-free summer Arctic Ocean by the mid-2030s with an 80% probability.

"Optimal Combination of Arctic Sea Ice Extent Measures: A Dynamic Factor Modeling Approach," *International Journal of Forecasting*, 2021

- Proposed and estimated a dynamic factor model that combines four measures of Arctic sea ice extent to account for differing volatility and cross-correlations.
- Used the Kalman smoother to extract an optimal combined measure.
- Confirmed and enhanced the confidence in the NSIDC Sea Ice Index.

"Measuring International Uncertainty: the Case of Korea," *Economics Letters*, 2018.

- Leverage a data rich environment to construct and study a measure of macroeconomic uncertainty for the Korean economy.
- Provide several stylized facts about uncertainty in Korea.
- Compared the proposed measure with financial and policy uncertainty proxies and found that neither financial nor policy uncertainty proxies capture economy-wide uncertainty.

WORKING PAPERS

"The Promise of Time-Series Foundation Models for Agricultural Forecasting: Evidence from Marketing Year Average Prices"

- Evaluated 17 forecasting approaches across econometrics, machine learning, deep learning, and foundation models (e.g., Chronos2, TimesFM 2.5, Time-MoE) on USDA monthly commodity price data.
- Demonstrated that zero-shot foundation models consistently outperform traditional econometrics, machine learning, and deep learning benchmarks.
- Showed that Time-MoE achieved $\sim$**45%** accuracy improvement over USDA futures-based forecasts, establishing a new paradigm for agricultural forecasting.

"Incorporating Prior Knowledge of Latent Group Structure in Panel Data Models."

- Introduced group patterns into heterogeneous coefficients of panel data models.
- Develop a constrained Bayesian grouped estimator that exploits researchers' prior beliefs on groups in a form of pairwise constraints.

- Provided better density forecasts for the disaggregated U.S. CPI inflation rate than standard time series and panel data models.
- Generated meaningful group structures in a cross-countries application.

"Bayesian Estimation of Sparsely Heterogeneous Panel Models"

- Developed a dynamic panel data model incorporating a spike-and-slab prior to address coefficient heterogeneity, capturing both sparse and dense heterogeneity in real data.
- Implemented Bayesian framework for accurate inference, leveraging Gibbs sampling for posterior inference under sparse heterogeneity conditions.
- Applied the model to income data from the PSID, revealing evidence of sparse heterogeneity in balanced panels with long-term employment histories.

WORK IN PROGRESS

"Estimate Nonlinear VARs using Neural Network Models"

- Proposed a simulation-based method to generate impulse response functions based on VAR models in daily frequency.
- Implement neural networks model to model the nonlinearity across macroeconomics variables.

"Early Warning Indicator of Banking Crises via Factor-Augmented VAR Models"

- Constructed macroeconomics and financial uncertainty indices for developed and developing countries as leading indicators to financial crises.
- Evaluated the accuracy of early warning indicators as an operational indicator to keep track of the business cycle.