Ke Xu

Research internships | Summer 2027

Ph.D. Candidate, University of Notre Dame | Expected graduation: May 2028

kxu6@nd.edu | Personal website | Google Scholar | GitHub

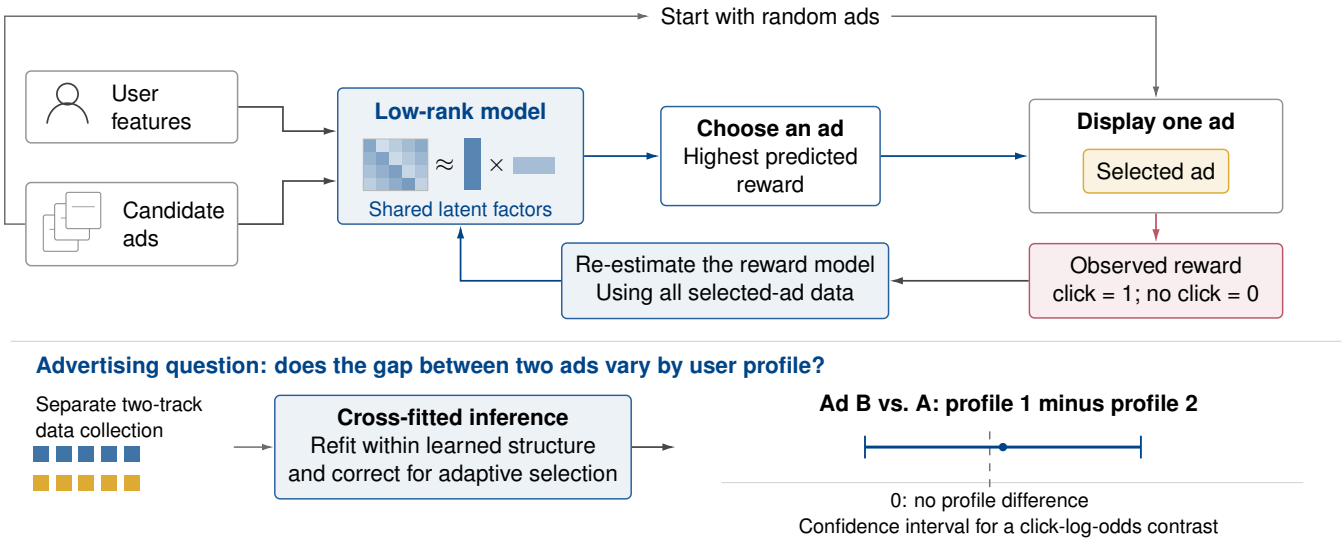
Making complex data useful for decisions

I enjoy finding structure in data that improves **estimation accuracy and computational efficiency** while enabling **reliable uncertainty quantification**. My research spans dependent time series, adaptively collected data, noisy tensor data, and causal inference from summary statistics. I develop methods to estimate underlying patterns and quantify the uncertainty in comparisons that inform decisions.

Learning and inference from adaptive feedback

How can we improve choices and learn reliably from feedback?

An advertising system observes responses only to the ads it shows, so its choices shape the evidence available for learning and inference. My two-stage policy learns **shared user–item structure**, then selects greedily without continued randomized exploration. It achieves **provably sublinear regret**, with diminishing average reward loss relative to the best available choices. Exploiting shared latent factors reduces the effective dimension governing the regret bound from the product to the sum of user and item feature dimensions, at fixed rank. I also developed a **novel cross-fitted inference method** for low-rank reward models that provides **asymptotically valid confidence intervals** for prespecified interaction comparisons under greedy data collection.



Selected work: Contextual bandits (working paper with Zifeng Zhao and Yuefeng Han). The separate two-track inference design targets prespecified linear contrasts at fixed rank. Ads and user profiles in the illustrative comparison are specified in advance.

Efficient estimation and inference for multiway data

How can we fit complex models efficiently and quantify uncertainty?

CP and Tucker tensor models exploit low-rank structure. For CP decomposition and scalar- and tensor-on-tensor regression, I developed Riemannian gradient descent (RGD) and Riemannian Gauss–Newton (RGN). Both have **convergence guarantees** to the statistical noise floor, with linear convergence for RGD and a quadratic convergence phase for RGN. Numerical experiments showed stable RGD convergence with strongly aligned latent factors. Both methods ran **faster with lower recovery error** than CP-ALS, the standard alternating-least-squares baseline for CP decomposition. Both are implemented in SegreCP [R/Rcpp package]. This work was accepted to ***NeurIPS*** 2026 (main-track poster).

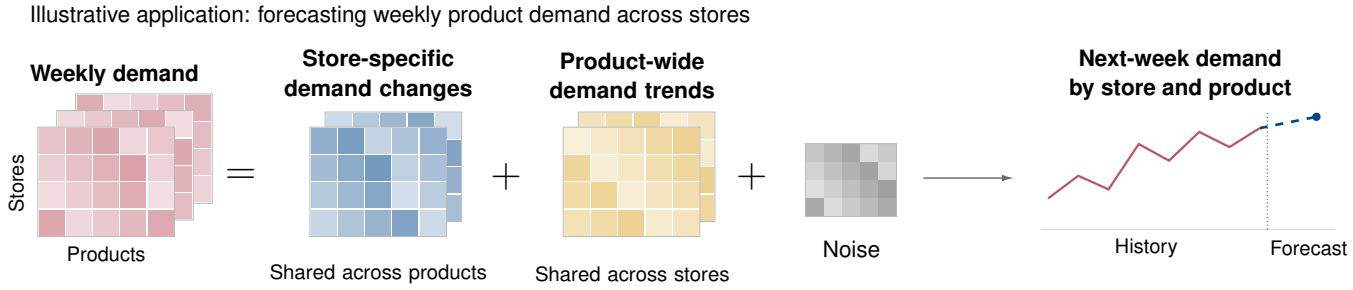
For Tucker models, I developed data-driven, asymptotically valid confidence intervals for entries, aggregate summaries, and contrasts in tensor regression and PCA, with lengths matching **minimax lower-bound scales**.

Selected work: Tensor optimization [earlier preprint] | Statistical inference for low-rank tensor models [preprint].

Modeling shared and local dynamics

How can forecasts capture both local demand and product-wide trends?

Shared trends can coexist with distinct local dynamics: store demand can change locally while a product's popularity shifts across locations. I co-developed a **modewise additive factor model** that represents these two sources of variation with **separate time-varying factors**, rather than forcing both through one shared factor core. Each component remains low-rank, sharing information across products or stores while preserving their distinct dynamics. Our estimation procedure **reduces interference between the components**, with recovery guarantees and uncertainty quantification for aligned factor loadings. For inventory planning, this offers a way to forecast store–product demand while preserving both local changes and product-wide trends.



Schematic: two dynamic low-rank components + noise. Business data and forecast are illustrative.

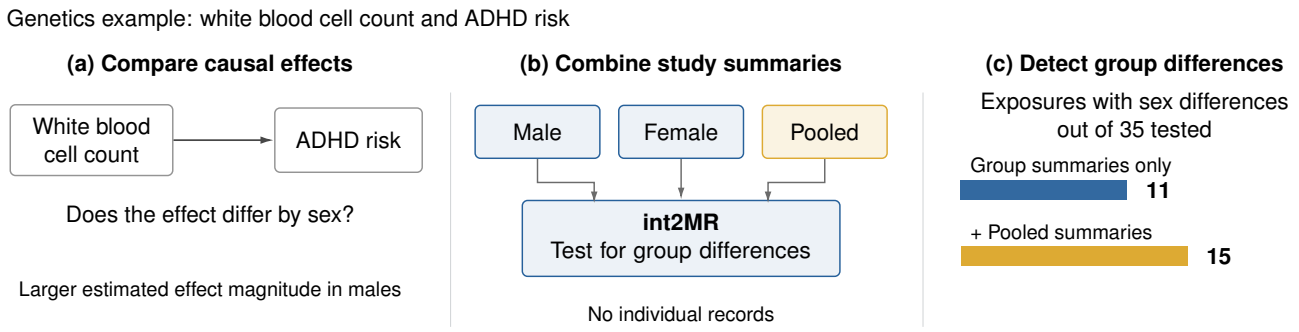
Evaluated on real data. My Python expanding-window pipeline used quarterly macroeconomic data from 18 OECD countries. The model had the lowest mean one-step-ahead error against four factor-model baselines across all reported evaluation windows.

Selected work: Modewise Additive Factor Model for Matrix Time Series [arXiv:2512.25025]. Under revision at ***Annals of Statistics***.

Causal inference from summary data

Can we compare causal effects without individual records?

An overall causal effect can hide group differences that matter for decisions. I developed **int2MR**, a genetic instrumental-variable method and R/Rcpp package that estimates and tests **group-specific causal effects without individual-level records**. By combining group-specific and pooled study summaries, it enables interaction testing and increases power while **maintaining Type I error control** in simulations. A related industry question is whether a product change increases conversion differently for new and returning users when only group-level results from randomized experiments can be shared.



Genetic instrumental-variable analysis of 35 exposures for ADHD, with a common false discovery rate threshold of 0.10. Bars summarize all exposures; white blood cell count is one example.

My collaborative work also uses **data integration to improve causal-effect estimation**, combining summary statistics across related outcomes and datasets.

Selected work: ***PLOS Genetics***, 2025 [paper] | int2MR [R/Rcpp package]
The American Journal of Human Genetics, 2024 [paper] | ***The American Journal of Human Genetics***, 2026 [paper].