October 2025

Kan Yao

Address: Department of Economics
Yale University
New Haven, CT 06520-8268

Telephone: +1 (818) 964-7128

E-mail: kan.yao@yale.edu

Web page: https://www.kan-yao.com

Fields of Concentration:
Primary Field: Econometrics
Secondary Field: Industrial Organization

Comprehensive Examinations Completed:
2020 (Oral): Econometrics and Industrial Organization
2019 (Written): Microeconomics and Macroeconomics

Dissertation Title: *Essays in Econometrics*

Committee:
Professor Yuichi Kitamura (Chair)
Professor Donald Andrews
Professor Xiaohong Chen

Education:
Ph.D., Economics, Yale University, 2025 (expected)
M.Phil., Economics, Yale University, 2022
M.A., Economics, Yale University, 2020
B.A., Economics & B.S., Applied Mathematics, UCLA, 2018

Fellowships, Honors and Awards:
University Dissertation Fellowship, Yale University, 2024–2025
Doctoral Fellowship, Yale University, 2018-2025
Cowles Foundation Fellowship, Yale University, 2018-2024
Nathan Hale Associates Fellowship, Yale University, 2019-2020
Economics Research Fellowship, UCLA, 2017-2018

Teaching Experience:
Spring 2024, 2025 Teaching Assistant to Prof. Edward Vytlacil, Econometrics (UG)
Fall 2022, Teaching Assistant to Prof. Cormac O'Dea, Introductory Microeconomics (UG)
Spring 2022, Teaching Assistant to Prof. Yuichi Kitamura, Econometrics (UG)
Fall 2021, Teaching Assistant to Prof. Nick Ryan, Intro to Data & Econometrics (UG)

Spring 2021, Teaching Assistant to Prof. Donald Andrews, Econometrics III (G)
Fall 2020, Teaching Assistant to Prof. Costas Meghir, Intermediate Econometrics (UG)

Research Experience:

Research Assistant to Prof. Yuichi Kitamura, Yale University, 2019
Research Assistant to Prof. Rodrigo Pinto, UCLA, 2016-2018

Work Experience:

Economist Intern, Amazon, Summer 2025

Working Papers:

"Low-Rank Estimation of Nonlinear Panel Data Models", *Job Market Paper*

Work In Progress:

"GMM Estimation of High-Dimensional Panels"
"Measuring Inflation Inequality with Incomplete Prices" with Olivia Ding, Tianyu Fan

Languages:

English (fluent), Chinese (native)

References:

Prof. Yuichi Kitamura
Yale University
Department of Economics
New Haven, CT 06520
Phone: 203-432-3699
yuichi.kitamura@yale.edu

Prof. Donald Andrews
Yale University
Department of Economics
New Haven, CT 06520
Phone: 203-432-3698
donald.andrews@yale.edu

Prof. Xiaohong Chen
Yale University
Department of Economics
New Haven, CT 06520
Phone: 203-432-5852
xiaohong.chen@yale.edu

Dissertation Abstract

Low-Rank Estimation of Nonlinear Panel Data Models [Job Market Paper]

Heterogeneity is central to economic modeling and empirical research, particularly in the analysis of panel data, where latent factors that vary over time and across individuals can influence outcomes and introduce bias if not properly addressed. For instance, in a demand model, purchase decisions depend not only on observed factors such as prices or income, but also on unobserved determinants such as tastes or risk tolerances. These unobserved factors can evolve over time, and consumers may also differ in how strongly they adjust their demand in response to observed variables like price. This motivates a flexible yet parsimonious framework that incorporates dynamic unobservables and also accommodates heterogeneous responses to observables in panel data.

I use interactive fixed effects (IFEs) to capture heterogeneous responses to unobserved time-varying shocks, while allowing these shocks to correlate arbitrarily with observables. Combining IFEs with heterogeneous responses to observables, however, makes estimation more complex and can create non-convexities in the parameters of interest such as those in random-coefficients models. These challenges highlight the need for methods that are both grounded in asymptotic theory and practical to implement. To address these challenges, I develop a unified framework for nonlinear panel models with IFEs, aiming to (i) deliver consistent and asymptotically normal parameter estimates and (ii) develop algorithms that are computationally efficient in large panels.

My proposed solution is a two-step estimation method. The first step applies nuclear-norm regularization (NNR), which exploits the low-rank structure of the interactive fixed effects to obtain consistent initial parameter estimates. Here, I implement this using a novel Majorization-Minimization algorithm. In the second step, I iteratively refine these estimates to improve convergence rates and achieve asymptotic normality. Since incidental parameter bias may still arise, I also discuss bias correction methods to address this issue.

This two-step approach is computationally feasible and applicable to a broad range of complex economic models. I conduct a simulation study to analyze the finite-sample performance of our method and assess the adequacy of our computational procedure for practical implementation. Additionally, I apply the method in an empirical setting to study the cross-market arbitrage behavior of nonfinancial firms.

GMM Estimation of High-Dimensional Panels

In this paper, I develop a general framework for high-dimensional panel data models characterized by moment restrictions and interactive fixed effects. The framework allows both the number of parameters and the number of moment conditions to grow with the sample size, encompassing a broad class of linear and nonlinear specifications. I focus on settings with many covariates, among which only a small subset has a nonzero impact on the parameters of interest, and unobserved heterogeneity is captured by a low-rank component structured through latent factors and their loadings. This structure poses challenges for conventional sparse estimators such as Lasso, and for standard factor estimators like PCA. To address these issues, I propose a new generalized method of moments (GMM) procedure that combines ℓ_1 and nuclear-norm penalties. Under general conditions, I show that the proposed estimator consistently recovers both the nonzero coefficients and the latent low-rank matrix in the presence of many moments. The findings contribute to the econometric literature by providing a practical solution for high-dimensional panel data models, ensuring both theoretical consistency and computational feasibility.