Yihong Liu
Research Statement

University of California, Santa Barbara

October 27th, 2026

My research is in theoretical and applied econometrics, with a focus on valid and robust causal inference in microeconomic settings. I seek to improve empirical practice by developing methods that are grounded in theory and straightforward to implement. My work is motivated by methodological challenges in applied research and by real-world economic problems.

My current research spans two areas: difference-in-differences (DiD) and triple-differences (DDD) designs, which are my primary focus, and structural estimation of dynamic games. In two working papers, including my job market paper, I examine commonly used DiD and DDD procedures when identification relies on conditioning on covariates. I identify limitations of these procedures and develop alternative methods that provide consistent estimation and valid inference. My third working paper develops a tractable generalized method of moments (GMM) framework for estimating structural parameters in dynamic games with nonlinear dynamics. In what follows, I will first summarize my completed research and then describe my work in progress and future directions.

Completed Research

Matched Triple-Differences: A Framework for Covariate Adjustment (Job Market Paper)

In my job market paper, I study covariate adjustment in DDD designs. A common empirical strategy is to include either covariate trends or covariate levels in a three-way fixed effects (3WFE) regression. This practice is typically motivated by a conditional parallel gaps assumption which assumes that the deviation from the parallel trends is similar across units with comparable observed covariates. My paper makes two main contributions.

First, I formally analyze both 3WFE specifications and show that in general, neither consistently estimates the average treatment effect on the treated (ATT). My analysis has two parts. I begin by showing that the OLS estimands of both specifications reweight the outcomes incorrectly and therefore fail to satisfy the covariate balancing condition that is sufficient for recovering the ATT. I then characterize the additional assumptions under which each OLS estimand equals the ATT. These assumptions impose strong restrictions on treatment effect heterogeneity, conditional outcome means, or the propensity score.

Second, I develop a matched triple-differences framework that addresses these limitations and provides valid estimation and inference. I propose a two-step procedure that combines covariate matching with a 3WFE regression to obtain a matched triple-differences estimator. I then establish the consistency and asymptotic normality of the estimator under a set of novel regularity conditions on the matching step. Because these conditions do not depend on any particular matching procedure, they yield a unified asymptotic theory for a broad class of matched triple-differences estimators. Notably, this class includes estimators based on commonly used nearest-neighbor matching, with either a fixed or a diverging number of neighbors, as well as kernel matching. To account for the uncertainty introduced by the matching step, I also propose a novel variance estimator that is consistent for every matching procedure covered by the framework. My empirical application shows that the proposed matched triple-differences estimator can yield estimates and qualitative conclusions that differ from those obtained using the 3WFE regressions.

Post-Matching Two-Way Fixed Effects Estimation

In joint work with Gonzalo Vazquez-Bare, we analyze the common practice of combining matching with two-way fixed effects (2WFE) methods. Researchers first match each treated unit to one or more never-treated units based on observed pre-treatment covariates. They then estimate treatment effects using a 2WFE regression in the matched sample, weighting never-treated units by the number of times they are matched.

We formally analyze this practice and highlight three problems. First, with variation in treatment timing, the post-matching 2WFE estimator that pools all treated cohorts has an asymptotic "pooling" bias, even when the treatment effect is constant across units and over time. Second, failing to account for the variability introduced by the matching step yields invalid standard error estimators, which can be biased upward or downward depending on the data generating process. Third, imperfect matches generate a matching bias that can be first-order for inference, but the required bias correction to address it is commonly omitted when using post-matching 2WFE estimators in empirical work. Thus, conventional t-statistics and confidence intervals based on the uncorrected estimator may be invalid.

To address these problems, we propose bias-corrected matched DiD estimators of treatment effects, event-study coefficients and other aggregate magnitudes, and we show that they are consistent and asymptotically normal. We also provide consistent standard errors that account for the matching step. We illustrate our results using simulations and two empirical applications.

Estimating Resource Games with Nonlinear Dynamics

Estimating structural parameters in dynamic resource games is challenging for two reasons: their equations of motion contain meaningful nonlinearities, such as biological reproduction or exponential disease transmission, and they involve strategic interactions among multiple players. In joint work with Shuchen Tsao, we develop a tractable GMM framework that addresses both challenges, allowing for heterogeneous agents and fully nonlinear, interactive state dynamics. We provide sufficient conditions under which the Markov-perfect Nash equilibrium admits analytical optimality conditions, and establish the consistency and asymptotic normality of the resulting GMM estimator. Applying the framework to the Japanese anchovy fishery shared by China, Japan, and Korea, we recover each country's implicit ecological valuation and simulate the equilibrium effects of climate-driven shifts in fish migration patterns. We further show that linearizing the equations of motion, the standard route to tractability, substantially understates these ecological values.

Work in Progress and Future Directions

Building on these projects, I plan to extend my work on DiD and DDD designs and to develop a complementary research agenda on partial identification of treatment effects in stratified and clustered experiments. I have work in progress and concrete research proposals in both areas.

The first direction continues my research on DiD and DDD designs. I plan to extend the matched triple-differences framework in my job market paper to settings with staggered adoption and repeated cross-sectional data. More broadly, I plan to investigate open methodological questions that arise from the empirical use of DiD and DDD designs.

One such question motivates my first work in progress "**Triple-Differences Designs with a Continuous Treatment**". Many applied microeconomic studies use DDD designs with a continuous treatment and estimate causal effects using 3WFE regressions. To my knowledge, however, formal methodological guidance for these designs remains limited. In this project, starting from a two-period setting, I will formally define the causal parameters of interest and establish the identification conditions. I will then examine whether commonly

used 3WFE specifications recover these causal parameters and point out the potential limitations. Building on this analysis, I will develop valid nonparametric estimation and inference procedures based on series or spline estimators. Finally, I will extend the core results to settings with instrumental variables, with covariates, and with multiple periods and variation in treatment timing.

A second direction concerns partial identification of treatment effects in stratified and clustered experiments. My second work in progress, "**Identification and Inference on Treatment Effects under Covariate-Adaptive Randomization and Sample Selection**," studies treatment effect bounds in randomized experiments with sample selection when treatment assignment depends on pre-treatment covariates. My current results have two implications. First, sharp bounds such as Lee and Manski bounds take a similar form under covariate-adaptive randomization as in the i.i.d. setting, but establishing their sharpness requires a separate argument because observations need not be i.i.d. under this design. Second, the inference for the bound estimators must account for the experimental design. Standard errors that ignore the design overstate the true sampling variability, leading to overly conservative inference. As this project develops, I plan to extend the identification, estimation, and inference results to experiments with multiple treatments and to develop a unified framework that also covers other stratified designs, such as matched-pair experiments.

This work in progress also provides a starting point for studying treatment effect bounds in stratified cluster-randomized experiments, where treatment is assigned at the cluster level but sample selection, such as attrition, occurs at the individual level. In these settings, applied researchers may target different estimands, such as individual-average and cluster-average treatment effects. However, bounds developed for experiments without clustering may neither target these estimands nor account for the dependence induced by cluster-level assignment. I plan to establish identification results and develop estimation and inference procedures that account for both the estimand of interest and the experimental design.

Looking further ahead, I intend to expand my research agenda to artificial intelligence (AI) and large language models (LLMs), which are rapidly becoming part of the empirical toolkit. In collaboration with other researchers, I plan to bring rigorous statistical inference to these tools, including inference for explainable AI methods and for the use of LLMs in non-parametric regression. As in my work on DiD and DDD designs, my goal is to provide theory-driven guidance that allows applied researchers to use these tools reliably and to draw credible conclusions from their data.