Statistics Winter Workshop: 2024

Causal Inference and its Applications

Friday, January 29th

9:00AM—5:00PM

J. Wayne Reitz Union Auditorium

Saturday, January 20th

9:00AM—12:30PM

Harn Museum Auditorium

Venues

Friday, January 19th

J. Wayne Reitz Union
655 Reitz Union Drive
University of Florida Campus
Gainesville, FL 32611

Rooms
Presentations: Auditorium
Refreshments: Grand Salon A & E
Posters: Room 3315

Parking

The Visitor Welcome Center and Bookstore parking garage is located at the Reitz Union, at the corner of Museum Road and Reitz Union Drive. The garage is an unrestricted pay facility available to all members of the university community. This garage can accommodate 300 vehicles. There are 45 short-term parking spaces located in the garage.

The garage hours of operation are Monday through Friday, 7:30 am to 4:30 pm. Short-term and daily fees apply during this time. The garage may be used during non-operating hours for short-term parking, free of charge.

Saturday, January 20th

Harn Museum
3259 Hull Road
Gainesville, FL 32608

Room
Harn Museum Auditorium

Parking

Parking on evenings and weekends is free.

Program—Day One

Friday, January 19, 2024

08:30 AM: **Breakfast (Grand Salons A & E)**

09:00 AM   **Welcome—Dr. Mike Daniels**

Professor and Chair, Department of Statistics, University of Florida

09:15 AM   **Dr. Damon Clark**

Associate Professor, Department of Economics, University of California-Irvine

"The Impacts of Attending Popular Schools: Quasi-Experimental Evidence from Lotteries and Equity Priorities", with Paul Yoo (UC Irvine) and Paco Martorell (UC Davis)"

10:00 AM   **Dr. Laura Giuliano**

Professor of Economics, University of California, Santa Cruz

"Can Gifted Education Help Higher-Ability Boys from Disadvantaged Backgrounds?"

10:45 AM   **Break (Grand Salons A & E)**

11:00 AM   **Dr. Kosuke Imai**

Professor of Government and Statistics, Harvard University

``Statistical Performance Guarantee for Subgroup Identification with Generic Machine Learning."

11:45 AM   **Lunch (Grand Salons A & E)**

01:00 PM   **Dr. Fabrizia Mealli**

Professor, Department of Economics, European University Institute

"Do Test Scores Help Teachers Give Better Track Advice to Students? A Principal Stratification Analysis"

01:45 PM   **Dr. Shu Yang**

Associate Professor of Statistics, North Carolina State University

"Test-Then-Pool: A uniformly valid inferential framework for data integration."

02:30 PM	Break (Room 3315, 3rd Floor)
02:45 PM	Dr. Matt Masten

Associate Professor, Department of Economics, Duke University

"Assessing the Importance of Omitted Variables in Linear Regression"

03:30 PM	Poster Session (Room 3315, 3rd Floor)

Program—Day Two

Saturday January 20, 2024

09:00 AM:	Breakfast (Auditorium)
09:30 AM:	Dr. Mark Hoekstra

Professor of Economics, Baylor University

"Are American Juries Racially Discriminatory? Evidence from more than a quarter million grand jury cases."

10:15 AM	Dr. Celeste Carruthers

Distinguished Professor, Haslam College of Business, University of Tennessee

"Free community college and college completion: Evidence from Tennessee"

11:00 AM	Break

11:15 AM	Dr. Dylan Small

Professor of Statistics, Department of Statistics and Data Science, University of Pennsylvania

"Exploratory Data Analysis, Confirmatory Data Analysis and Replication in the Same Observational Study: A Two Team Cross-Screening Approach to Studying the Effect of Unwanted Pregnancy on Mothers' Later Life Outcomes

Dr. Damon Clark

Associate Professor, Department of Economics, University of California-Irvine

Title: The Impacts of Attending Popular Schools: Quasi-Experimental Evidence from Lotteries and Equity Priorities", with Paul Yoo (UC Irvine) and Paco Martorell (UC Davis)

Abstract: Many school districts operate "school choice" or "open enrollment" programs that give parents a choice of school. The popular schools in these districts are often oversubscribed, such that districts must decide which applicants receive priority at these schools. While many school districts prioritize by lottery, a growing number give priority to disadvantaged students. Using variation in popular school attendance generated by an equity priority and lotteries, this paper investigates whether - and which - students benefit from attending popular schools.

Biography: Damon Clark is an Associate Professor (with tenure) in the Department of Economics at the University of California, Irvine. He is also a Research Associate at the NBER, a Research Fellow at IZA and a Research Fellow at the Institute for Fiscal Studies. His is Associate Editor at the American Economic Journal: Applied Economics, the Journal of the European Economic Association and the Journal of Human Resources. His main research interests are in the economics of education, labor economics and public economics.

Dr. Laura Giuliano

Professor of Economics, University of California, Santa Cruz

Title: Can Gifted Education Help Higher-Ability Boys from Disadvantaged Backgrounds?

Abstract: Children from disadvantaged backgrounds are less likely than their peers to succeed in school and enter college. Disparities are wider for boys, leading to large college enrollment gaps between low-income girls and boys of similar cognitive ability. We study how gifted status in 5th grade – determined by having IQ≥116 – affects subsequent educational outcomes of disadvantaged children in a large urban school district. For marginally eligible boys, gifted identification raises the college entry rate by 25-30 percentage points, closing the gap with their female counterparts. Analyses of course-taking and grades point to a key role for engagement in math.

Biography: Laura Giuliano is a Professor of Economics at the University of California, Santa Cruz and a Research Associate at the National Bureau of Economic Research. She has also held faculty positions at UC Merced and the University of Miami and visiting positions at UC Berkeley and the University of Virginia. In 2015–16 she served as Senior Economist for Labor, Education and Welfare for President Obama's Council of Economic Advisers. Giuliano's research addresses topics within the broad fields of labor economics and education by applying quasi-experimental methods to non-experimental data. Her research in education includes impact evaluations of gifted and advanced academic programs and of policies that increase participation in these programs by students from underserved groups; it has been published in such outlets as the American Economic Review and the Proceedings of the National Academy of Sciences.

Dr. Kosuke Imai

Professor of Government and Statistics, Harvard University

Title: Statistical Performance Guarantee for Subgroup Identification with Generic Machine Learning.'

Abstract: Across a wide array of disciplines, many researchers use machine learning (ML) algorithms to identify a subgroup of individuals who are likely to benefit from a treatment the most (``exceptional responders'') or those who are harmed by it. A common approach to this subgroup identification problem consists of two steps. First, researchers estimate the conditional average treatment effect (CATE) using an ML algorithm. Next, they use the estimated CATE to select those individuals who are predicted to be most affected by the treatment, either positively or negatively. Unfortunately, CATE estimates are often biased and noisy. In addition, utilizing the same data to both identify a subgroup and estimate its group average treatment effect results in a multiple testing problem. To address these challenges, we develop uniform confidence bands for estimation of the group average treatment effect sorted by generic ML algorithm (GATES). Using these uniform confidence bands, researchers can identify, with a statistical guarantee, a subgroup whose GATES exceeds a certain effect size, regardless of how this effect size is chosen. The validity of the proposed methodology depends solely on randomization of treatment and random sampling of units. Importantly, our method does not require modeling assumptions and avoids a computationally intensive resampling procedure. A simulation study shows that the proposed uniform confidence bands are reasonably informative and have an appropriate empirical coverage even when the sample size is as small as 100. We analyze a clinical trial of late-stage prostate cancer and find a relatively large proportion of exceptional responders. The open-source software package is available for implementing the proposed methodology.

Biography: Kosuke Imai is Professor in the Department of Government and the Department of Statistics at Harvard University. He is also an affiliate of the Institute for Quantitative Social Science. Imai specializes in the development of statistical methods and machine learning algorithms and their applications to social science research. His areas of expertise include causal inference, computational social science, and survey methodology.

Dr. Fabrizia Mealli

Professor, Department of Economics, European University Institute

Title: Do Test Scores Help Teachers Give Better Track Advice to Students? A Principal Stratification Analysis

Abstract: Every year more than one million EU students receive high-school track recommendations from their teachers. There is no clear evidence that teachers are "good" at providing these recommendations. There is not even a consensus on what constitutes a "good" recommendation. Our goal is to establish, using a quasi-randomized experiment and a well-defined metric based on Principal Stratification, whether teachers can give "better" recommendations with or without the information provided by standardized test scores. We find that, on average, test scores improve the capacity of teachers to give better track advice to students. Joint work with Gustav Axén, Andrea Ichino, and Javier Viviens Martín

Biography: Fabrizia Mealli is Professor of Econometrics at the Department of Economics of the European University Institute (EUI) since September 2023. She is on leave as Professor of Statistics at the Department of Statistics, Informatics, Applications "G. Parenti" - DiSIA of the University of Florence. Her research focuses on causal inference, program evaluation, estimation techniques, simulation methods, missing data, and Bayesian inference, with applications to the economics, social and biomedical sciences. She is an associate editor of serval journals which include Observational Studies, The Annals of Applied Statistics, Journal of the American Statistical Association, T&M, and Biometrika.

Dr. Shu Yang

Associate Professor of Statistics, North Carolina State University

Title: Test-Then-Pool: A uniformly valid inferential framework for data integration

Abstract: This talk addresses the challenges of pre-testing estimation, a prevalent yet complex practice in statistical analysis due to the double usage of data. I will illustrate these challenges through a "Test-Then-Pool" analysis used in data integration, introducing a data-adaptive method for selecting the optimal test statistic threshold. This method minimizes the mean square error of the estimator for our target parameter. I will demonstrate the non-regularity of the Test-Then-Pool analysis using the local alternative device, highlighting the failure of standard inference procedures such as bootstrap, sample splitting, and soft thresholding. Lastly, I will present an adaptive confidence interval that ensures uniform validity and exhibits strong performance in finite samples.

Biography: Shu Yang is Associate Professor of Statistics, Goodnight Early Career Innovator, and University Faculty Scholar at North Carolina State University. She received her Ph.D. in Applied Mathematics and Statistics from Iowa State University and postdoctoral training at Harvard T.H. Chan School of Public Health. Her primary research interest is causal inference and data integration, particularly with applications to comparative effectiveness research in health studies. She also works extensively on methods for missing data and spatial statistics. She has been Principal Investigator for several U.S. NSF, NIH, and FDA research projects.

Dr. Matt Masten

Associate Professor, Department of Economics, Duke University

Title: Assessing the Importance of Omitted Variables in Linear Regression

Abstract: Omitted variables are one of the most important threats to the identification of causal effects. In this paper we survey recent methods for assessing the impact of omitted variables on empirical conclusions, with a particular focus on methods that allow researchers to calibrate sensitivity parameters by comparing the magnitude of selection on observables with the magnitude of selection on unobservables. We illustrate these approaches in two empirical applications to the effect of historical American frontier life on modern cultural beliefs and the relationship between social capital and the rise of the Nazi party. We implement these methods in the companion State module regsensitivity for easy use in practice."

Biography: Matt is an Associate Professor in the Department of Economics at Duke University. I am an econometrician working on identification and causal inference. His current focus is on robustness and sensitivity analysis.

Dr. Mark Hoekstra

Professor of Economics, Baylor University

Title: Are American Juries Racially Discriminatory? Evidence from more than a quarter million grand jury cases

Abstract: To what extent are the large racial disparities in the criminal justice system caused by taste-based or statistical discrimination based on race? We examine this question using over a quarter million felony cases from Texas in which grand juries decided whether or not the case should be prosecuted. We estimate disparate impact—defined as whether juries treat Black defendants differently from White defendants with the same underlying level of felony guilt—by exploiting the quasi-random assignment of cases to grand juries to purge omitted variable bias. Results from cases involving defendants with racially identifiable names indicate a small, but statistically significant, disparate impact of 0.8 percent against Black defendants. In order to distinguish racial bias from alternative sources of disparate impact, and to address concerns that racial bias later in the process could impact the conviction outcomes used for our test, we also compare Black and White defendants with similarly-White-sounding names. These defendants were thus racially indistinguishable to grand jurors, but not to prosecutors or trial juries later in the process. Results indicate a disparate impact of the same magnitude. This indicates that jurors do not engage in statistical or taste-based discrimination.

Biography: Mark Hoekstra is the George J. Boden Professor of Economics at Baylor University. Professor Hoekstra's research in empirical microeconomics addresses a wide range of questions relevant for public policy. His previous work has focused on education, law and economics, and elections. Currently he is focused on understanding how and why neighborhoods matter for child and adult outcomes, and whether outcomes in the criminal justice system are shaped by race or gender.

Dr. Celeste Carruthers

Distinguished Professor, Haslam College of Business, University of Tennessee

Title: Free community college and college completion: Evidence from Tennessee

Abstract: Free community college" programs clarify and reduce the cost of college. These programs can increase college enrollment, but one concern is that they will route students away from four-year institutions and delay or hinder bachelor's degree attainment. We study the staggered rollout of tuition-free community college in Tennessee, which culminated with statewide Tennessee Promise and preceded close to 30 similar programs in other states as well as two federal proposals. We find that access to free community college in Tennessee increased college enrollment by 2 - 6 percentage points for 18 – 19 year-olds (3 – 9% of the mean). By age 22, eligible students were 2 percentage points (26%) more likely to have attained an associate's degree and imprecisely more likely to have attained a bachelor's degree.

Biography: Celeste Carruthers is the William F. Fox Distinguished Professor of Labor Economics in the Haslam College of Business at the University of Tennessee with a joint appointment in the Department of Economics and the Boyd Center for Business and Economic Research. Her research centers on education policy with crossovers into public economics, labor economics, and economic history. Recent and ongoing projects examine the effect of financial aid on college choices, career and technical education, and the consequences of segregated schools in the early 20th-century United States. She teaches graduate and undergraduate courses in public expenditure analysis, causal inference, and econometrics. Carruthers is the Editor-in-Chief of Economics of Education Review, a Research Associate at the National Bureau of Economic Research, a member of the CTE Research Network at the American Institutes for Research, a Baker Expert at the Howard H. Baker Jr. Center for Public Policy, and a former faculty advisor to several fellows in the Harvard Graduate School of Education Strategic Data Projec

Dr. Dylan Small

Professor of Statistics, Department of Statistics and Data Science, University of Pennsylvania

Title: Exploratory Data Analysis, Confirmatory Data Analysis and Replication in the Same Observational Study: A Two Team Cross-Screening Approach to Studying the Effect of Unwanted Pregnancy on Mothers' Later Life Outcomes

Abstract: Exploratory data analysis, confirmatory data analysis and replication are three important aspects of building strong evidence from observational studies. Exploratory data analysis, confirmatory data analysis and replication are often thought of as being done on separate studies. However, for settings where randomized experiments are impossible to conduct for ethical reasons and observational studies must be relied on, it is common that there is a data set with unique strengths. We develop a two-team cross screening approach that allows for exploratory data analysis, confirmatory data analysis and replication to be done in the same observational study data set. We apply the approach to study the effect of unwanted pregnancy on mothers' later life outcomes using data from the Wisconsin Longitudinal Study. This is joint work with Samrat Roy, Marina Bogomolov, Ruth Heller, Amy Claridge and Tishra Beeson.

Biography: Dylan Small is the Universal Furniture Professor of Statistics in the Department of Statistics and Data Science, The Wharton School, University of Pennsylvania. His research focuses on causal inference and its applications to health and the social sciences.