European Causal Inference Meeting 2021, 2nd July

Mixed themes: Instrumental variables, experimental designs, econometrics and applications

This meeting is hosted by the LSE Department of Statistics and is a free Zoom event

Time	Session/talk information
10:00-10:10	Welcome and Introduction **Sara Geneletti**, LSE, UK
10:10-12:35	**Session1: Missing data/Quasi-experimental designs** **Chair: Chengchun Shi**
10:10-10:35	**Manuel Gomes**, UCL, UK *Addressing missing data in the estimation of time-varying treatments*
10:35-11:00	**Marica Valente**, ETH Zurich, Switzerland/ DIW Berlin, Germany *Heterogenous effects of waste pricing policies*
11:00-11:25	**Fiammetta Menchetti**, University of Florence, Italy *Estimating the causal effect of an intervention in a time series setting: the C-ARIMA approach*
11:25-11:35	**10 min Break**
11:35-12:35	**Chair: Sara Geneletti**, LSE, UK **Keynote Speaker Ingeborg Waernbaum**, University of Uppsala, Sweden *Calibration/entropy balancing estimators for average causal effects – model implication and robustness properties*
12:35-13:35	**Lunch**
13:35-14:35	**Session2: Instrumental variables/Mendelian Randomization** **Chair: Joshua Öoftus**
13:35-14:00	**William Denault**, Center of excellence for fertility and health, Norway *Cross-fitted instrument: a v´blueprint for one-sample Mendelian Randomization*
14:00-14:25	**Carlos Cinelli**, UCLA, USA *An Omitted Varibale Bias Frameowrk for Sensitivity Analysis of Instrumental Variables*
14:25-14:35	**10 min Break**
14:35-15:24	**Flash talks: 7 minute mini-presentations** **Chair: Sara Geneletti**, LSE, UK
14:35-14:42	**Jack Bowden**, University of Ester, UK *The Triangulation Within a STudy (TWIST) framework for causal inference within Pharmac-ogenetics research*
14:42-14:49	**Imke Mayer**, Ecole des Hautes Sciences Sociale, France *Transporting treatment effects with incomplete attributes*
14:49-14:56	**Chan Park**, University of Wisconsin – Madison, USA *Assumption-Lean Analysis of Cluster Randomized Trials in Infectious Diseases for Intent-to-Treat Effects and Spillover Effects Among A Vulnerable Subpopulation*
14:56-15:03	**Aaron Sarvet**, Harvard University, UK *Non-parametic inference for counterfactual population parameters under time-varying*
15:03-15:10	**Kancharla Manjusha**, Univerity of Wisconsin – Madison, USA *Robust Randomized Experiments for Causal Effects Under Privacy*
15:10-15:17	**Zach Branson**, Carnegie Mellon University, USA *Randomization Tests for Assessing Covariate Balance When Designing and Analyzing Matched Datasets*
15:17-15:24	**Emmett Kendall**, North Carolina State University; University of Florida, USA *Testing local treatment effects using regression discontinuity designs that are robust to a violation of causal assumptions*
15:25-15:55	**Discussion of Flash Talks**
15:55-16:00	Thanks and closing remarks **Sara Geneletti**, LSE, UK

Abstracts

Manuel Gomes, Juan Sgura-Buisan
Addressing missing data in the estimation of time-varying treatments

Routinely-collected data are increasingly used to establish the effectiveness of health interventions, particularly for estimating treatment strategies sustained over time, i.e. time-varying treatments. Inverse probability weighting (IPW)-based marginal structural models (MSMs) are commonly used to address the time-varying confounding in these studies. IPW can also be used to tackle any missing data in the outcomes and/or confounders, but this often leads to biased, inefficient estimates of treatment effects. Previous studies have combined multiple imputation (MI) with MSMs for addressing missing data in studies with time-varying confounding but focused on missing confounders and mostly monotone missingness. This paper compares MI and IPW to tackle both missing outcomes and confounders, across both monotone and intermittent missing data settings in the evaluation of time-varying treatments.
This study is motivated by an evaluation of sustained biologic treatment of patients with severe rheumatoid arthritis, using data from the US National Data Bank for rheumatic diseases. About 25% of participants had missing data in the outcomes (health-related quality-of-life and costs) or confounders. Addressing the missing data with MI led to somewhat different treatments effects and narrower confidence intervals compared to IPW.
Through a comprehensive simulation study, we assessed the performance of the proposed approach across different scenarios with monotone and non-monotone missingness in the outcome and confounders, and alternative mechanisms and proportions of missing data. We found that combining MI with MSMs provided the lowest biases and root mean squared error compared to standard IPW-based MSMs across the wide range of scenarios considered in this study.

Marica Valente
Heterogeneous effects of waste pricing policies

Using machine learning methods in a quasi-experimental setting, I study the heterogeneous effects of introducing waste prices - unit prices on household unsorted waste disposal - on waste demands and social welfare. First, using a unique panel of Italian municipalities with large variation in prices and observables, I show that waste demands are nonlinear. I find evidence of nudge effects at low prices, and increasing elasticities at high prices driven by income effects and waste habits before policy. Second, I estimate policy impacts on pollution and municipal management costs, and compute the overall social cost savings for each municipality. Social welfare effects become positive for most municipalities after three years of adoption, when waste prices cause significant waste avoidance.

Fiamette Menchetti, Iavor Bojinov
Estimating the causal effect of an intervention in a time series setting: the C-ARIMA approach

The Rubin Causal Model (RCM) has a long tradition in the definition and the estimation of causal effects in the context of randomized experiments and, in recent years, several methods have been developed to estimate the effects of interventions occurring in time series settings. However, none of these makes use of ARIMA models, which are instead very common in the econometrics literature. Bridging the gap between causal inference and econometrics, we propose a novel approach, C-ARIMA, to define and estimate the causal effect of an intervention in a time series setting under the RCM. We first formalize the assumptions enabling the definition, the estimation, and the attribution of the causal effect to the intervention; we then describe three causal estimands and their estimators. The latter are defined by comparison of the observed response with a predicted counterfactual in the absence of intervention, as resulting from a forecast of the ARIMA model fitted in the pre-intervention period. To perform inference, we derive three hypothesis tests for the causal effect estimators. We then check the validity of the proposed method with an extensive simulation study, comparing its performance against a standard intervention analysis approach. In the empirical application, we use C-ARIMA to assess the causal effect on supermarket sales of a permanent price reduction on selected store-brand products.

Ingeborg Waernbaum
Properties of calibration estimators of the average causal effect – A comparative study of balancing properties

Causal analyses with observational data require adjustment for confounding variables. Properties of semi-parametric estimators using estimated propensity scores, conditional outcomes and a combination thereof

with different degrees of flexibility of parametric models have been in focus in the causal literature in recent years. Early guidance to model selection suggested that model specification, fitting and balance checking could be performed in an iterative procedure. This was followed by proposals of, now standard, doubly robust AIPW estimators that fit parametric models for the propensity score and conditional outcomes given covariates. More recently, a class of weighting estimators have been proposed that directly aim at incorporating covariate balance in the estimation process through calibration/entropy maximization.
Since covariate balance is not a sufficient condition for identification of the true propensity score the general calibration estimator, using finite constraints, has an asymptotic error which depends on the covariance of the error of an implicit propensity score fit and the conditional outcomes. Although here, as for the AIPW estimators, robustness properties are implicit in the estimation procedure. In this talk we describe weighting estimators within the more recent calibration/entropy balancing proposals (Tan, 2020, Chan et al. 2016) and other alternatives to propensity score estimation such as a kernel approach, RKHS (Wong and Chan, 2018) and the covariate balancing propensity score, CBPS (Imai and Ratkovic 2014, Fan et al. 2018). We describe and compare asymptotic properties for calibration/entropy balancing estimators using Kullback-Leibler and quadratic Rényi divergence (Källberg and Waernbaum, 2020) with the related logit calibration estimator proposed by Tan and the CBPS estimator by Imai and co-authors. The estimators are applied to data from the Swedish Childhood Diabetes Register in a study of the effect of school achievements on complications Type 1 Diabetes Mellitus. The finite-sample properties of the estimators are investigated in a simulation study and compared to their corresponding asymptotic errors. The simulations also include an evaluation of variance estimators proposed for the calibration estimators in Källberg and Waernbaum (2020), Tan (2020) and Chan et al. (2016). This is joint work with David Källberg and Emma Persson, Umeå University.

William Denault, Jon Bohlin, Stephen Burges, Astanand Jugessur
Cross-fitted instrument: a blueprint for one-sample Mendelian Randomization

Bias from weak instruments may undermine the ability to estimate causal effects in instrumental variable regression (IVR). We present here a simple solution for handling strong confounding and weak instruments bias by introducing a new type of instrumental variable, coined 'cross-fitted instrument' (CFI). Our approach based on CFI entails partitioning the data at random and estimating the impact of the instrument on the exposure in each partition, and then using these estimates to perform an IVR on each partition. We adapt CFI to Mendelian Randomization (MR) and termed this adaptation 'cross-fitting for Mendelian Randomization' (CFMR). CFMR uses all the available data to select the genetic instruments and estimates the effect of the exposure on the outcome, thus avoiding the use of a large amount of the data to only selecting the instruments, as usually performed in two-sample MR. Therefore, CFMR enhances the power of MR in the context of a meta-analysis, as it allows performing an unbiased one-sample MR in each cohort and then meta-analyzing the results across the cohorts without the need to use any cohort to select the genetic instruments. Moreover, CFMR allows performing cross-ethnic MR while accounting for heterogeneity due to ethnicity, which is important in consortium-led meta-analyses where the cohorts might be of different ethnicity. To our knowledge, no MR approach is currently can account for such heterogeneity. Finally, CFMR enables the application of MR to exposures that are rare or difficult to measure, which would normally preclude their analysis in the regular two-sample MR setting.

Carlos Cinelli, Chad Hazlett
An Omitted Variable Bias Framework for Sensitivity Analysis of Instrumental Variables

We develop an “omitted variable bias” framework for sensitivity analysis of instrumental variable (IV) estimates that is immune to “weak instruments,” naturally handles multiple “side-effects” (violations of the exclusion restriction assumption) and “confounders” (violations of the ignorability of the instrument assumption), exploits expert knowledge to bound sensitivity parameters, and can be easily implemented with standard software. Conveniently, we show that many pivotal conclusions regarding the sensitivity of the IV estimate (e.g. tests against the null hypothesis of zero causal effect) can be reached simply through separate sensitivity analyses of two familiar auxiliary OLS estimates, namely, the effect of the instrument on the treatment (the “first stage”) and the effect of the instrument on the outcome (the “reduced form”). More specifically, we introduce sensitivity statistics for routine reporting, such as robustness values for IV estimates, describing the minimum strength that omitted variables need to have to invalidate the conclusions of an IV study. Next we provide visual displays that fully characterize the sensitivity of IV point-estimates and confidence intervals to violations of the standard IV assumptions. Finally, we offer formal bounds on the worst possible bias under the assumption that the maximum explanatory power of omitted variables are no stronger

than a multiple of the explanatory power of observed variables. We apply our methods in a running example that uses instrumental variables to estimate the returns to schooling.

Jack Bowden, Luke C. Pilling, Deniz Türkmen, Chia-Ling Kuo, David Melzer
The Triangulation WIthin A STudy (TWIST) framework for causal inference within Pharmaco-genetics research

Over the last 20 years the field of Epidemiology has embraced the exploitation of random genetic inheritance to help uncover causal mechanisms of disease using the technique of Mendelian randomization (MR). Genetic variants can also play an important role in helping to explain treatment effect heterogeneity, through the science of pharmacogenetics. A canonical example is Clopidogrel: the primary drug for stroke prevention in the UK and many other countries. It requires CYP2C19 enzyme activation in order to be properly metabolised and thus work to its fullest extent.

We review the methodological underpinnings of the general pharmaco-genetic approach, which utilises only individuals who are treated and relies on fairly strong baseline assumptions to estimate what we refer to as the "genetically mediated treatment effect" (GMTE). When these assumptions are seriously violated, we show that a robust estimate of the GMTE that incorporates information on the population of untreated individuals can instead be used. In cases of partial violation, we clarify when Mendelian randomization and a modified confounder adjustment method can also yield consistent estimates for the GMTE. A full decision framework is then described to decide when a particular estimation strategy is most appropriate and how estimates can be combined to improve efficiency. We illustrate these approaches by re-analysing UK Biobank-CPRD linked data relating to CYP2C19 genetic variants, Clopidogrel use and stroke risk, and data relating to ApoE genetic variants, statin use and Coronary Artery Disease. We then discuss how the framework can be extended for use in other epidemiological contexts.

Imke Mayer, Julie Josse; Traumabase Group
Transporting treatment effects with incomplete attributes

The simultaneous availability of experimental and observational data to estimate a treatment effect is both an opportunity and a statistical challenge: Combining the information gathered from both data is a promising avenue to build upon the internal validity of randomized controlled trials (RCTs) and a greater external validity of observational data, but it raises methodological issues, especially due to different sampling designs inducing distributional shifts. We focus on the aim of transporting a causal effect estimated on an RCT onto a target population described by a set of covariates. Available methods such as inverse propensity weighting are not designed to handle missing values, which are however common in both data. In addition to coupling the assumptions for causal identifiability and for the missing values mechanism and to defining appropriate strategies, one has to consider the specific structure of the data with two sources and treatment and outcome only available in the RCT.
We study different approaches and their underlying assumptions on the data generating processes and distribution of missing values and suggest several adapted methods, in particular multiple imputation strategies.
These methods are assessed in an extensive simulation study and practical guidelines are provided for different scenarios.
This work is motivated by the analysis of a large registry of over 20,000 major trauma patients and a multi-centered RCT studying the effect of tranexamic acid administration on mortality.

Chan Park, Hyunseung Kang
Assumption-Lean Analysis of Cluster Randomized Trials in Infectious Diseases for Intent-to-Treat Effects and Spillover Effects Among A Vulnerable Subpopulation

Cluster randomized trials (CRTs) are a popular design to study the effect of interventions in infectious disease settings. However, standard analysis of CRTs primarily relies on strong parametric methods, usually a Normal mixed effect models to account for the clustering structure, and focus on the overall intent-to-treat (ITT) effect to evaluate effectiveness. The paper presents two methods to analyze two types of effects in CRTs, the overall and heterogeneous ITT effects and the spillover effect among never-takers who cannot or refuse to take the intervention.

For the ITT effects, we make a modest extension of an existing method where we do not impose parametric models or asymptotic restrictions on cluster size. For the spillover effect among never-takers, we propose a

new bound-based method that uses pre-treatment covariates, classification algorithms, and a linear program to obtain sharp bounds. A key feature of our method is that the bounds can become dramatically narrower as the classification algorithm improves and the method may also be useful for studies of partial identification with instrumental variables. We conclude by reanalyzing a CRT studying the effect of face masks and hand sanitizers on transmission of 2008 interpandemic influenza in Hong Kong.

Aaron Sarvet, Jessica Young, Kerolles N. Wanis, James M. Robins, Mats J. Stensrud
Non-parametric inference for counterfactual population parameters under time-varying allocation of scarce binary treatments

Emerging scarcity requires new policies for triaging limited resources. In such settings, sensible strategies often delay - rather than prevent - treatment reception for some patients, thus necessitating a time-varying regime. Additionally, when data arise from setting where treatment resources are newly limited, as in a crisis, standard assumptions on the independence of units are not satisfied. We formulate a general potential-outcomes-based framework for evaluating the effects of strategies for allocating a fixed supply of limited resources in a time-varying setting. We provide non-parametric conditions that allow identification and consistent estimation of counterfactual population parameters under such regimes for finite and large target populations using only a single draw from a finite-population of causally-connected units. As an illustration, we consider estimation of survival under counterfactual rules for ventilator triage (including both initiation and termination) in an intensive care unit over the course a COVID-19 epidemic. We show that triage rules that optimize for short-term survival may have sub-optimal survival by the end of a mass-casualty event in many settings.

Kancharla Manjusha, Hyunseung Kang
Robust Randomized Experiments for Causal Effects Under Privacy

Randomized control trials (RCTs) have been the gold standard to estimate the average causal effect of a treatment on a response. However, many randomized experiments assume that participants are willing to share their personal data, specifically their response to treatment. This assumption, while trivial at first, is becoming difficult to satisfy in the modern era where there are more regulations to protect users' data and privacy. The paper presents a simple experimental design that is differentially private, one of the strongest notions of data privacy in computer science. Simply put, the data collected from our design is information-theoretically guaranteed to be truly anonymous; the same guarantee does not exist for a traditional RCT where an adversary can potentially identify individual patients based on data from the RCT. Critically, our design's strong guarantee on data privacy enables individual-level data to be publicly shared for scientific replication. We also use works on non-compliance in experimental psychology to make our design robust against "adversarial" study participants who may distrust the investigator with their personal data and provide contaminated responses to intentionally bias the results of the study. Under our new design, we propose unbiased and asymptotically Normal estimators for the average treatment effect. We also present a doubly robust estimator that leverages pre-treatment covariates, if available. We apply our design to two settings where data privacy is important: a RCT evaluating therapies for mental health and a RCT evaluating different modes of learning in online statistics courses at the University of Wisconsin-Madison.

Zach Branson
Randomization Tests for Assessing Covariate Balance When Designing and Analyzing Matched Datasets

Causal analyses for observational studies are often complicated by covariate imbalances among treatment groups, and matching methodologies alleviate this complication by finding subsets of treatment groups that exhibit covariate balance. It is widely agreed upon that covariate balance can serve as evidence that a matched dataset approximates a randomized experiment, but what kind of experiment does it approximate? We develop a randomization test for the hypothesis that a matched dataset approximates a particular experimental design, such as complete randomization, block randomization, or rerandomization. Our test can incorporate any experimental design and allows for a graphical display that puts several designs on the same univariate scale, thereby allowing researchers to pinpoint which design--if any--is most appropriate for a matched dataset. After researchers determine a plausible design, we recommend a randomization-based analytical approach, which can incorporate any design and treatment effect estimator. Through simulation, we find that our test can frequently detect violations of randomized assignment that harm inferential results. Furthermore, through a real application in political science, we find that matched datasets with high levels of covariate balance tend to approximate balance-constrained designs like rerandomization, and analyzing them

as such can lead to precise causal analyses. However, assuming a precise design should be proceeded with caution, because it can harm inferential results if there are still substantial biases due to remaining imbalances after matching. Although we focus on matching, we also demonstrate how to use this approach to assess covariate balance in instrumental variable analyses and regression discontinuity designs.

Emmett Kendall, Brenden Beck, Joseph Antonelli
Testing local treatment effects using regression discontinuity designs that are robust to a violation of causal assumptions

We study the degree of variation in policing outcomes that is attributable to differential policing practices in New York City (NYC) using geographic regression discontinuity designs. By focusing on small geographic windows near precinct boundaries we can estimate local average treatment effects of police precincts. The geographic regression discontinuity design relies heavily on continuity assumptions of the potential outcomes near the boundary of interest. While these assumptions are often thought to be more realistic than other assumptions used to infer causality from observational data, they can easily be violated in realistic applications. We develop a novel and robust approach to testing whether there are differences in policing outcomes that are caused by differences in police precincts across NYC. In particular, our test is robust to violations of the assumptions traditionally made in geographic regression discontinuity designs, and we show that valid inference can be obtained under much weaker assumptions. We utilize a unique form of resampling to identify new geographic boundaries that are known to have no treatment effect, which provides a valid estimate of the estimator's null distribution even under violations of standard assumptions. We find that this procedure gives drastically different results in the analysis of NYC arrest rates than those that rely on standard assumptions, thereby providing more robust estimates of the nature of the effect of police precincts on arrest rates in NYC.