Symposium on Machine Learning for Causal Inference in the Health and Social Sciences

Date: 15th of December 2023

Time: 9am to 16:30 pm

Venue: London Mathematical Society (de Morgan House, Hardy room)
https://maps.app.goo.gl/ZjiznQwp6nUUsHm16

Financial and other support: Medical Research Council, Wellcome Trust, Centre for Data and Statistical Sciences for Health (DASH) at the London School of Hygiene and Tropical Medicine, UCL Department of Economics, UCL Department of Statistical Science, Centre for Microdata Methods and Practice (Cemmap) and the Institute for Fiscal Studies

Time	Session	Details
9:00- 9:20	Registration and refreshments	
9:20- 9:30	Welcome	
9:30- 10:30	Session 1: Statistics	• Stijn Vansteelandt (University of Ghent) - IV-learner: learning conditional average treatment effects using instrumental variables • Karla DiazOrdaz (UCL Department of Statistical Science) - Non-parametric variable importance measures for heterogeneous causal effects • Oliver Hines (QuantCo) - Causal inference with continuous treatments, a tale of two estimands Chair: Aureo de Paula (UCL, Department of Economics, IFS)
10:30- 10:45	Break and refreshments	
10:45- 11:45	Session 2: Economics	• Martin Weidner (University of Oxford, Department of Economics) - A Neyman Orthogonalization Approach to the Incidental Parameter Problem • Julia Hatamyar (University of York, Centre for Health Economics) - Machine Learning for Difference-in-differences with Staggered Adoption and Dynamic Treatment Effect Heterogeneity • Liyang Sun (UCL, Department of Economics) - Empirical Welfare Maximization with Constraints Chair: Gianluca Baio (UCL, Department of Statistical Science)
11:45- 12:00	Break	
12:00- 13:00	Session 3: Health Economics & Health Data Science	• Noemi Kreif (University of York, Centre for Health Economics) - Policy Learning with Rare Outcomes • Stephen O'Neill (London School of Hygiene and Tropical Medicine) - An approach for combining clinical judgment with machine learning to inform medical decision-making • David Glynn (University of York, Centre for Health Economics) - Integrating decision modelling and machine learning Chair: Ioanna Manolopoulou (UCL, Department of Statistical Science)

13:00- 14:00	Lunch	
14:00- 15:00	Keynote lecture	• Whitney Newey (Massachusetts Institute of Technology) - TBA Chair: Karla DiazOrdaz (UCL, Department of Statistical Science)
15:00- 15:15	Break and refreshments	
15:15- 16:30	Panel discussion and closing remarks	Topic: Making (causal) machine learning useful for policy: model validation and other challenges Panellists: • Ioanna Manolopoulou (UCL, Department of Statistical Science) • Richard Grieve (London School of Hygiene and Tropical Medicine) • Chris Harbron (Roche) • Stephen Hansen (UCL, Department of Economics) Chair: Noemi Kreif (University of York, Centre for Health Economics)

Presentation abstracts

Presenter: Stijn Vansteelandt (University of Ghent)

Title: IV-learner: learning conditional average treatment effects using instrumental variables (based on joint work with Stephen O'Neill, Karla Diaz-Ordaz and Richard Grieve)

Abstract: Driven by clinical inquiries regarding the efficacy of emergency surgery for gastrointestinal conditions, I will discuss the estimation of conditional average treatment effects (CATE) in the presence of unmeasured confounding using instrumental variables. While data-adaptive methods (e.g. statistical learning or machine learning) offer relief from concerns related to model misspecification bias inherent in parametric approaches, they introduce their own challenges, notably regularization bias and potential overfitting. Specifically, slow convergence rates affecting the first stage (machine learning based) regression of exposure on instrument and covariates, may propagate into the CATE estimates, resulting in poor accuracy. Synthetic data simulations confirm this, but moreover reveal poor performance of existing strategies for constructing a so-called Neyman-orthogonal learner, which is nonetheless designed to mitigate regularization bias in first-stage predictions. To address this, I will propose an alternative Neyman-orthogonal learner by strategically tailoring first-stage predictions to excel in their ultimate task: delivering CATE estimates with low regularization bias. Simulation studies validate substantial enhancements in performance, underscoring the effectiveness of the proposed approach.

Presenter: Karla DiazOrdaz (UCL, Department of Statistical Science)

Title: Non-parametric variable importance measures for heterogeneous causal effects

(Based on joint work with Oliver Hines and Stijn Vansteelandt)

Abstract: There is a growing interest in methods that focus on finding “personalized” treatment rules given the measured characteristics individuals. Policy learning focuses on finding optimal treatment rules (OTRs) but does not provide insight into the extent to which treatment benefits (or harms) individual subjects. Techniques for quantifying treatment effect heterogeneity in terms of the conditional average treatment effects (CATEs) do offer such insights but may represent a complicated function of the individuals' characteristics that provide little insight into the key drivers of heterogeneity. Justifiably, decision makers and stakeholders are rightly hesitant to blindly adopt such treatment policies. To address these limitations, we introduce novel nonparametric treatment effect variable importance measures (TE-VIMs). TE-VIMs extend recent regression-VIMs, viewed as nonparametric analogues to ANOVA statistics. We derive estimators (based on their efficient influence curves) and illustrate their use through a simulation study and an applied example.

Presenter: Oliver Hines (QuantCo)

Title: Causal inference with continuous treatments, a tale of two estimands
(based on joint work with Karla DiazOrdaz and Stijn Vansteelandt)

Abstract: In economics and medicine one is often interested in the main effect of a continuous treatment (dose, price, duration) on an outcome. Dose-response curve modelling is a popular approach, but the uniform interventions considered may be unrealistic for some subpopulations. Worse still, estimators may be poorly supported by the data due to the extrapolation needed to predict expected outcomes under unrealistic treatments. This talk focuses on two emerging alternatives. Average Derivative Effects (ADEs) indicate the mean direction and magnitude of small changes to the treatment around the values observed in the data. Least Squares Estimands (LSEs) represent the “coefficient” when the outcome is projected (in a nonparametric sense) on to a partially linear model. In our work we showed that LSEs are optimally weighted ADEs and both belong to a common class that also contains the average treatment effect (a ubiquitous measure for the main effect of a binary treatment on an outcome). This talk gives some motivation and intuition behind these results/ estimands.

Presenter: Martin Weidner (University of Oxford, Department of Economics)

Title: A Neyman Orthogonalization Approach to the Incidental Parameter Problem
(based on joint work with Stephane Bonhomme and Koen Jochmans)

Abstract: A popular approach to perform inference on a parameter in the presence of nuisance parameters is to construct estimating equations that are orthogonal to those, in the sense that their expected first derivative is zero. However, usual first-order orthogonalization methods may fail when the nuisance parameters are poorly estimated. Leading economic examples of poorly estimated parameters are fixed effects in panel and network data. In this paper we focus on conditional likelihood models, and show how to perform higher-order orthogonalization and de-bias the estimator of interest up to any order. We show how high-order orthogonalization, combined with sample splitting, effectively reduces bias in a variety of applications. In panel data models, this provides a way to reduce the incidental parameter bias to any order. In an empirical application to a model of team production, we find that higher-order orthogonalization provides an effective way to reduce bias.

Presenter: Julia Hatamyar (University of York, Centre for Health Economics)

Title: Machine Learning for Difference-in-differences with Staggered Adoption and Dynamic Treatment Effect Heterogeneity
(based on joint work with Noemi Kreif, Martin Huber and Rudi Rocha)

Abstract: We combine two recently proposed nonparametric difference-in-differences methods, extending them to enable the examination of treatment effect heterogeneity in the staggered adoption setting using machine learning. The proposed method, machine learning difference-in-differences (MLDID), allows for estimation of time-varying conditional average treatment effects on the treated, which can be used to conduct detailed inference on drivers of treatment effect heterogeneity. We perform simulations to evaluate the performance of MLDID and find that it accurately identifies the true predictors of treatment effect heterogeneity. We then use MLDID to evaluate the heterogeneous impacts of Brazil's Family Health Strategy on infant mortality.

Presenter: Liyang Sun (UCL, Department of Economics)

Title: Empirical Welfare Maximization with Constraints

Abstract: When designing eligibility criteria for welfare programs, policymakers naturally want to target the individuals who will benefit the most given their cost to the program. This paper extends the previous literature on Empirical Welfare Maximization (EWM) for selecting eligibility criteria based on data by allowing for uncertainty in estimating the budget needed to implement the criterion, in addition to its benefit. Due to the additional estimation error, the EWM rule no longer selects eligibility criteria that consistently achieve the highest benefit possible while satisfying a budget constraint uniformly. The lack of uniformity is shown to apply to any statistical rule. I also propose two new statistical rules that perform better than the EWM rule under a budget constraint, and use them to select eligibility criteria for Medicaid expansion based on experimental data, a setting with imperfect take-up and varying program costs.

Presenter: Noemi Kreif (University of York, Centre for Health Economics)

Title: Policy Learning With Rare Outcomes

(based on joint work with Julia Hatamyar)

Abstract: Machine learning (ML) estimates of conditional average treatment effects (CATE) can guide policy decisions, either by allowing targeting of individuals with beneficial CATE estimates, or as inputs to decision trees that optimise overall outcomes. There is limited information available regarding how well these algorithms perform in real-world policy evaluation scenario, including those with rare outcomes such as infant mortality. Using synthetic data, we compare the finite sample performance of different policy learning algorithms, machine learning techniques employed during their learning phases, and methods for presenting estimated policy values. For each algorithm, we assess the resulting treatment allocation by measuring deviation from the ideal ("oracle") policy. Our main finding is that policy trees based on estimated CATEs outperform trees learned from doubly-robust scores. Across settings, Causal Forests and the Normalised Double-Robust Learner perform consistently well, while Bayesian Additive Regression Trees perform poorly. These methods are then applied to a case study targeting optimal allocation of subsidised health insurance, with the goal of reducing infant mortality in Indonesia.

Presenter: Stephen O'Neill (London School of Hygiene and Tropical Medicine)

Title: An approach for combining clinical judgment with machine learning to inform medical decision-making

(based on joint work with S. Moler-Zapata, A. Hutchings, R. Grieve, R. Hinchliffe, N. Smart, S. R. Moonesinghe, G. Bellingan, R. Vohra and S. Moug)

Abstract: Background Machine Learning (ML) methods can identify complex patterns of treatment effect heterogeneity. However, before ML can help to personalise decision-making, transparent approaches must be developed that draw on clinical judgement. We develop an approach that combines clinical judgment with ML to generate appropriate comparative effectiveness evidence for informing decision-making. We motivate this approach in evaluating the effectiveness of 'non-emergency surgery' (NES) strategies, such as antibiotic therapy, compared to emergency surgery (ES) for people with multiple long-term conditions (MLTCs) who have acute appendicitis. Our four-stage approach: draws on clinical judgement about which patient characteristics and LTCs modify the relative effectiveness of NES (i), selects additional covariates from a high-dimensional covariate space (P>500) by applying ML methods to large-scale administrative data (N=24,312) (ii), generates estimates of comparative effectiveness for relevant subgroups (iii), and presents evidence in a suitable form for decision-making (iv). This approach provides useful evidence for clinically-relevant subgroups. We found that overall the NES strategy led to increases in the mean number of days alive and out-of-hospital compared to ES, but estimates differed across subgroups, ranging from 21.2 (95% CI: 1.8 to 40.5) for patients with chronic heart failure and chronic kidney disease, to -10.4 (-29.8 to 9.1) for patients with cancer and hypertension. Our interactive tool for visualising ML output allows findings to be customised according to the specific needs of the clinical decision-maker. This principled approach of combining clinical judgement with ML models can improve trust, relevance and usefulness of the evidence generated for clinical decision-making.

Presenter: David Glynn (University of York, Centre for Health Economics)

Title: Integrating decision modelling and machine learning

(based on joint work with Noemi Kreif, Julia Hatamyar, Ankur Pandya, John Giardina and Marta Soares)

Abstract: There is increasing interest in moving away from "one size fits all" approaches towards stratifying treatment decisions. Understanding how expected effectiveness and cost- effectiveness varies with patient covariates is a key aspect of stratified decision making. Recently proposed machine learning (ML) methods can learn heterogeneity in outcomes without pre-specifying subgroups or functional forms, enabling the construction of decision rules ("optimal policies") that map individual covariates into a treatment decision. However, these methods do not yet integrate ML estimates into a decision modelling framework to reflect long-term policy-relevant outcomes and synthesise information from multiple sources. In this paper, we propose a method to integrate ML and decision modelling, when individual patient data is available to estimate treatment-specific survival time. We also propose a novel implementation of optimal policy algorithms to define subgroups using decision model output. We demonstrate these methods using the SPRINT (Systolic Blood Pressure Intervention Trial), comparing outcomes for "standard" and "intensive" blood pressure targets. We find the highest expected incremental net health benefit when treatment choice is allowed to be based on a complex function of covariates, followed by a tree-based algorithm to define subgroups.