Programme: Encounters in Econometric Theory
Nuffield College, May 13-14, 2022

Christophe Gaillac and Claudia Noack

Programme

Friday

10.00 - 10.30	Welcome
10.30 - 11.15	Sami Stouli
11.15 - 11.30	Coffee break
11.30 - 12.45	Jad Beyhum & Xiyu Jiao
12.45 - 14.00	Lunch
14.00 - 15.15	Kenichi Nagasawa & Claudia Noack
15.15 - 15.45	Coffee break
15.45 - 17.15	Kirill Borusyak & Ao Wang
17.15 - 17.45	Coffee Break
17.45 - 18.30	Andrei Zeleneev
18.45 - 21.30	Dinner

Saturday

08.30 - 09.15	Breakfast
09.30 - 10.45	Vishal Kamat & Christophe Gaillac
10.45 - 11.15	Coffee Break
11.15 - 12.45	Cristina Gualdani & Mengshan Xu
12.45 - 14.00	Lunch
at 14.00	Departure

Speakers and Talks

External Speakers

1. Jad Beyhum, CREST

Instrumental variable estimation of dynamic treatment effects on a survival outcome

This paper considers identification and estimation of the causal effect of the time Z until a subject is treated on a survival outcome T. The treatment is not randomly assigned, T is randomly right censored by a random variable C and the time to treatment Z is right censored by min(T,C). The endogeneity issue is treated using an instrumental variable explaining Z and independent of the error term of the model. We study identification in a fully nonparametric framework. We show that our specification generates an integral equation, of which the regression function of interest is a solution. We provide identification conditions that rely on this identification equation. For estimation purposes, we assume that the regression function follows a parametric model. We propose an estimation procedure and give conditions under which the estimator is asymptotically normal. The estimators exhibit good finite sample properties in simulations. Our methodology is applied to find evidence supporting the efficacy of a therapy for burn-out.

2. Kirill Borusyak, UCL

Non-Random Exposure to Exogenous Shocks

joint with Peter Hull

We develop a new approach to estimating the causal effects of treat- ments or instruments that combine multiple sources of variation ac- cording to a known formula. Examples include treatments capturing spillovers in social or transportation networks and simulated instru- ments for policy eligibility. We show how exogenous shocks to some, but not all, determinants of such variables can be leveraged while avoiding omitted variables bias. Our solution involves specifying coun- terfactual shocks that may as well have been realized and adjusting for a summary measure of non-randomness in shock exposure: the aver- age treatment (or instrument) across shock counterfactuals. We use this approach to address bias when estimating employment effects of market access growth from Chinese high-speed rail construction.

3. Cristina Gualdani, Queen Mary University

Price Competition and Endogenous Product Choice in Networks: Evidence from the US Airline Industry

joint with Christian Bontemps and Kevin Remmy

We build and estimate a two-stage model of airline competition in which firms first form the networks of markets to be served and then compete in prices. We estimate our model using fares data on US domestic flights. We show that large hub-and-spoke operations lower marginal costs but increase fixed costs. We evaluate a merger between American Airlines and US Airways and compare it to the bankruptcy and disappearance of American Airlines. We also evaluate remedies imposed by the Department of Justice on the merging parties and find evidence they limited harm to consumers.

4. Vishal Kamat, Toulouse School of Economics

Estimating Welfare Effects in a Nonparametric Choice Model: The Case of School Vouchers
joint with Samuel Norris
We develop new robust discrete choice tools to learn about the average willingness to pay and average cost of a school voucher in a program that randomly allocates vouchers. We consider a nonparametric, nonseparable choice model that places no restrictions on the functional form of utilities or the distribution of unobserved heterogeneity. We exploit the insight that the welfare parameters in this model can be expressed as functions of the demand for the different schools. However, while the random allocation of the voucher reveals the value of demand at two prices, the parameters generally depend on its values beyond these prices. We show how to sharply characterize what we can learn when demand is specified to be entirely nonparametric or to be parameterized in a flexible manner, both of which imply that the parameters are not necessarily point identified. We use our tools to analyze the welfare effects of voucher provision in the DC Opportunity Scholarship Program, a school voucher program in Washington, DC. We find that the provision of both the status-quo voucher and a wide range of counterfactual vouchers of different amounts have positive benefits net of costs. In comparison, traditional logit models produce estimates towards the lower end of our bounds, and hence may understate the benefits. We also find that the positive results can be explained by the popularity of low-tuition schools in the program; removing them from the program can result in a negative net benefit.

5. Kenichi Nagasawa, University of Warwick

Treatment effect estimation with noisy conditioning variables

In this paper, I develop a new identification strategy for treatment effects when observed variables are noisy measurements of unobserved confounding factors. I use proxy variables to construct a random variable conditional on which treatment variables become exogenous. The key idea is that, under appropriate conditions, there exists a one-to-one mapping between the distribution of unobserved confounding factors and the distribution of proxies. To ensure sufficient variation in the constructed control variable, I use an additional variable, termed excluded variable, which satisfies certain exclusion restrictions and relevance conditions. I establish asymptotic distributional results for flexible parametric and nonparametric estimators of the average structural function. I illustrate empirical relevance and usefulness of my results by estimating causal effects of college selectivity on wages.

6. Sami Stouli, University of Bristol

Gaussian Transforms Modeling and the Estimation of Distributional Regression Functions
joint with Richard Spady
Conditional distribution functions are important statistical objects for the analysis of a wide class of problems in econometrics and statistics. We propose flexible Gaussian representations for conditional distribution functions and give a concave likelihood formulation for their global estimation. We obtain solutions that satisfy the monotonicity property of conditional distribution functions, including under general misspecification and in finite samples. A Lasso-type penalized version of the corresponding maximum likelihood estimator is given that expands the scope of our estimation analysis to models with sparsity. Inference and estimation results for conditional distribution,

quantile and density functions implied by our representations are provided and illustrated with an empirical example and simulations.

7. Ao Wang, University of Warwick

Identification and (Fast) Estimation of Nonlinear Panel Models with Additively Separable Two-Way Fixed Effects
joint with Martin Mugnier
In this paper, we propose a class of nonlinear two-way fixed effect panel models that allow for unobserved heterogeneity in both intercept (additively separable with time fixed effect) and slopes (interacting with covariates) across individuals. The latter is particularly relevant when the researcher is interested in the distributional causal effects of covariates. We show that the fixed-effect parameters and the (nonparametrically specified) link function can be identified when T is large. To estimate the model, we propose a novel iterative Gauss-Seidel procedure. It is numerically equivalent to the MLE under standard conditions and overcomes the practical challenge of dimensionality in the number of fixed-effect parameters. Extensive Monte Carlo simulations suggest that the proposed method largely outperforms routine implementation of the MLE, especially when the number of fixed effects is large. Finally, we revisit two classic empirical studies in international trade (Helpman et al. 2008) and innovation (Aghion et al. 2013). Our empirical findings suggest non-negligible unobserved heterogeneity in the causal relationship of interest (trade elasticity and effect of institutional ownership on firm innovation, respectively) among countries and firms, confirming the empirical relevance of the proposed method.

8. Mengshan Xu, University of Mannheim

Policy choice in time series by empirical welfare maximization
joint with Toru Kitagawa and Weining Wang
This paper develops a novel method to inform policy choice in a dynamic setting where the available data is a single realization of multi-variate time-series. Building on the framework of statistical treatment choice, we propose Time-series Empirical Welfare Maximization (T-EWM) methods that estimate an optimal policy rule in the current period or over multiple periods by maximizing the empirical welfare crite- rion constructed upon nonparametric potential outcome time-series. We characterize conditions to consistently learn by T-EWM an optimal policy choice in terms of the conditional welfare given the history. We then derive a nonasymptotic upper bound of the conditional welfare regret and its minimax lower bound. To illustrate implemen- tation and uses of T-EWM, we perform simulation studies and apply the methods to estimate optimal monetary policy rules with macroeconomic time-series data.

9. Andrei Zeleneev, UCL

Robust Estimation and Inference in Panels with Interactive Fixed Effects.
joint with Timothy B. Armstrong and Martin Weidner
This paper considers linear panel regressions with unobserved factors and provides inference methods on the regression coefficients that are robust towards a violation of the "strong factor assumption" that is required in Bai (2009) and its follow-up literature. This is achieved by convexifying the low-rank constraint imposed by the factor model and afterwards applying general inference methods on linear models with convex constraints, see e.g. Donoho (1994) and Armstrong and Kolesar (2018). The resulting robust estima- tion and inference methods are shown to have good asymptotic properties and are very well-behaved at finite sample in our Monte Carlo simulations as well.

Internal Speakers

1. Christophe Gaillac, University of Oxford/Nuffield College

   Partially Linear Models under Data Combination
   joint with Xavier D'Haultfoeuille and Arnaud Maurel
   We consider the identification of and inference on a partially linear model, when the outcome of interest and some of the covariates are observed in two different datasets that cannot be linked. This type of data combination problem arises very frequently in empirical microeconomics. Using recent tools from optimal transport theory, we derive a constructive characterization of the sharp identified set. We then build on this result and develop a novel inference method that exploits the specific geometric properties of the identified set. Our method exhibits good performances in finite samples, while remaining very tractable. Finally, we apply our methodology to study intergenerational income mobility over the period 1850-1930 in the United States. Our method allows to relax the exclusion restrictions used in earlier work while delivering confidence regions that are informative.

2. Xiyu Jiao, University of Oxford/ Nuffield College.

   An Asymptotic Study of the False Outlier Detection Rate in Robust Two Stage Least Squares Models
   joint with Jonas Kurle
   A frequent concern in applied economics is that key empirical findings may be driven by a tiny set of outliers. To perform outlier robustness checks in practical applications of instrumental variables regressions, the common practice is first to run ordinary two stage least squares (2SLS) and remove observations classified as outliers with residuals beyond a chosen cut-off value. Subsequently 2SLS is re-calculated based on non-outlying observations, and this procedure can also be iterated until robust results are obtained. However, the above trimmed 2SLS has a positive probability of finding outliers even when the data generation process contains none. To answer the question whether observations identified are truly outliers, this paper studies the concept of false outlier detection rate (FODR) asymptotically using the empirical processes argument. With the established asymptotic theory of FODR, we propose two sets of tests for the overall presence of outliers. First, 'simple' tests on whether the proportion (or number) of detected outliers deviates from its expected value. Second, 'scaling' tests on whether the proportion (or number) of detected outliers decreases proportionally with the increasing level of the cut-off or heavy tail reference distribution used to detect outliers. Moreover, the FODR asymptotics provides a guidance for setting the cut-off value. Simulation studies lend further support to the theory and an empirical illustration to Acemoglu et al. (2019) shows the utility.

3. Claudia Noack, University of Oxford/Nuffield College

   Flexible Covariate Adjustments in Regression Discontinuity Designs
   joint with Tomasz Olma and Christoph Rothe
   Empirical regression discontinuity (RD) studies often use covariates to increase the precision of their estimates. In this paper, we propose a novel class of estimators that use such covariate information more efficiently than the linear adjustment estimators that are currently used widely in practice. Our approach can accommodate a possibly large number of either discrete or continuous covariates. It involves running a standard RD analysis with an appropriately modified outcome variable, which takes the form of the difference between the original outcome and a function of the covariates. We characterize the function that leads to the estimator with the smallest asymptotic variance, and

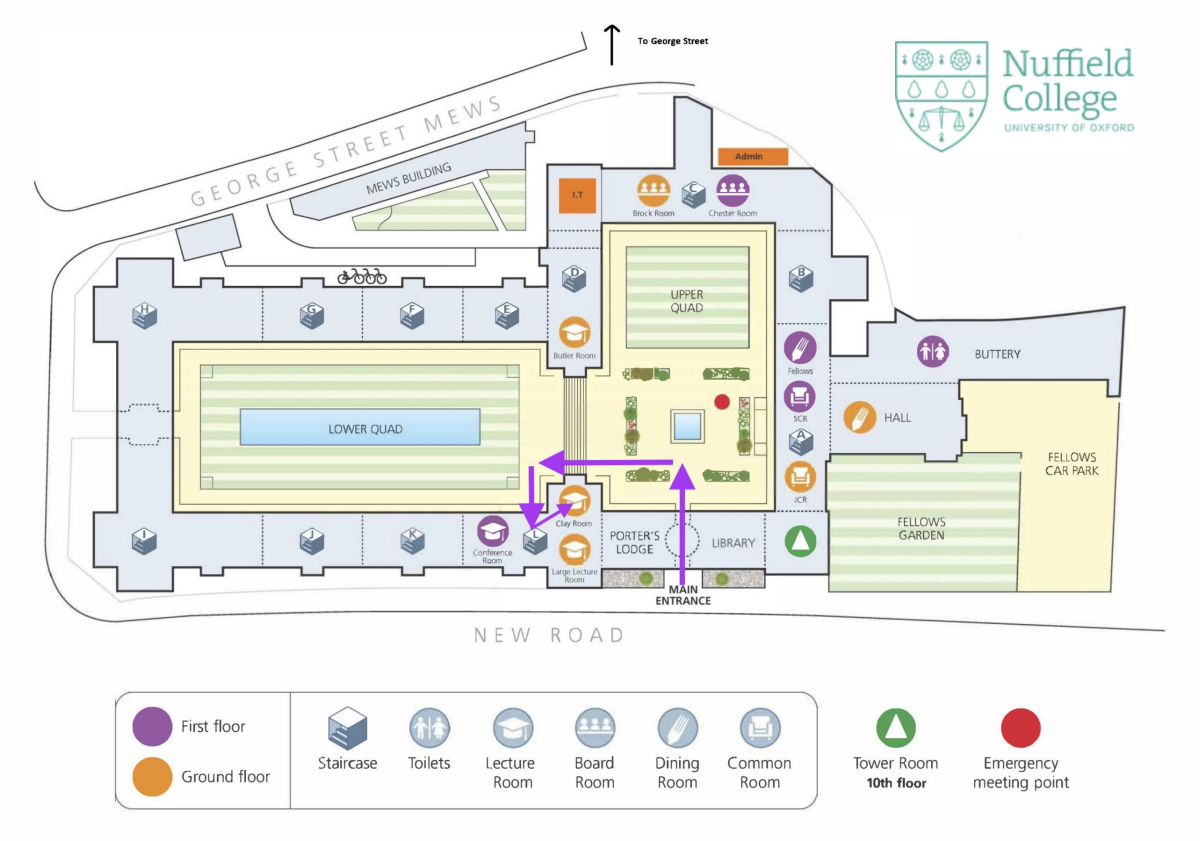
show how it can be estimated via modern machine learning, nonparametric regression, or classical parametric methods. The resulting estimator is easy to implement because tuning parameters can be chosen as in a conventional RD analysis.

Conference Venue

The presentations take place in the Clay Room:

Nuffield College
New Road
Oxford
OX1 1NF

An illustration of the conference venue:



We are very grateful to Nuffield College for hosting and fully financing the conference.