MUNI

Annex No. 10 to the MU Directive on Habilitation Procedures and Professor Appointment Procedures

HABILITATION THESIS REVIEWER'S REPORT

Masaryk University

Applicant	Lukáš Lafférs
Habilitation thesis	Essays in econometrics of model uncertainty
Reviewer	Prof. Anthony Strittmatter, Ph.D.
Reviewer's home unit, institution	Faculty of Economics, UniDistance Suisse

The thesis comprises two chapters. **Chapter 1** integrates the double machine learning (DML) framework into mediation analysis and dynamic treatment evaluation, addressing the challenges of very many confounding variables. It highlights DML as a solution to the biases introduced by standard machine learning estimators and illustrates its effectiveness using Neyman orthogonality principles. **Chapter 2** examines the challenges of point identification in econometric models, particularly the shortcomings of standard identification assumptions in non-experimental studies. It presents robust methods for computing bounds on causal effects. These bounds allow sensitivity analysis in the presence of model uncertainty or violation of identification assumptions. Each chapter has two subsections, which are detailed below:

- **Chapter 1.1** extends the application of DML to mediation analysis, addressing the challenges associated with very large numbers of confounding variables. The focus is on highlighting the limitations of standard ML estimators for causal inference and proposing DML as a remedy. By illustrating the principles of Neyman orthogonality, this chapter shows how DML effectively mitigates the biases introduced by ML estimators in the context of mediation analysis.

- **Chapter 1.2** further extends the application of DML, this time focusing on the evaluation of dynamic treatment effects in the context of sequential treatments. The chapter addresses the challenges of non-experimental treatment allocation and the shortcomings of naive comparisons. The study assumes sequential conditional independence. The key contribution is the introduction of a data-driven approach to accounting relevant covariates using DML, which effectively addresses the complexities arising from rich covariate information in the context of sequential treatments.

- **Chapter 2.1** focuses on the challenges associated with treatment/mediator endogeneity and outcome attrition. The chapter proposes a method for bounding direct and indirect effects under relaxations of identifying assumptions. The practical application of this methodology is demonstrated with a decomposition of the gender pay gap.

- **Chapter 2.2** discusses sensitivity analysis within sample selection models and presents a new computational method. It introduces an innovative way to relax identifying assumptions, in particular the exogeneity and monotonicity assumptions. The chapter proposes a new method for estimating bounds on average treatment effects when these assumptions are violated.

All chapters meet the level of scientific rigour expected in economics. They demonstrate the candidate's ability to develop and apply new econometric methods. Lukáš Laffér's personal contributions to the co-authored studies were conceptualisation, methodology, software development, data analysis and writing up of the paper. The quantity of contributions is proportional to the number of co-authors in each study, which is a common practice in economics.

Reviewer's questions for the habilitation thesis defence:

Chapter 1.1:

- Given the complexity of mediation analysis and the challenges posed by the large number of confounding variables, are there situations in which DML may still have limitations or provide suboptimal results (e.g. endogenous controls)?

- Are there situations in which the assumptions underlying Neyman orthogonality principles may not hold, potentially affecting the validity of the demonstrated efficiency of DML?

Chapter 1.2:

- Considering the assumption of sequential conditional independence, are there realistic situations where this assumption might not be plausible, and how would this affect the validity of the results?

- The sequential treatment approach appears to be very data intensive. What are the minimum data requirements to obtain meaningful results?

Chapter 2.1:

- Can you explain the advantage or difference between bounds and standard confidence intervals? What are the different concepts they capture?

- Why are bounds still not widely used in practice? Are there practical limitations or is the concept too complex for practitioners?

Chapter 2.2:

- Are there scenarios where the proposed method might fail to estimate meaningful bounds? For example, when sample selection is extreme.

Conclusion

The habilitation thesis entitled *Essays in econometrics of model uncertainty* by Lukáš Lafférs **fulfils** requirements expected of a habilitation thesis in the field of Economics.

Date: 27.01.2024   Signature: