REPORT ON MASTER THESIS
CENTER FOR ECONOMIC RESEARCH AND GRADUATE EDUCATION

STUDENT:	Aslan Bakirov
ADVISOR:	Paolo Zacchia
TITLE OF THE THESIS:	Revisiting Treatment Effects with Causal Forests

OVERALL ASSESSMENT *(provided in English, Czech, or Slovak):*

Please provide your assessment of each of the following categories, summary and suggested questions for the discussion. The minimum length of the report is 300 words.

The thesis consists of two, roughly equally long parts. First a summary of the methodological literature on causal effect estimation, with an emphasis on heterogeneous treatment effect estimation and machine learning techniques. Second, an application of the arguably most popular machine learning technique to uncover heterogeneity in treatment effects, Causal Random Forests (CRF), to the data of a recent paper on gender wage gaps, which did not attempt to uncover heterogeneity in the gaps (i.e. on the "causal effect of being female on wages"). Other papers have looked at heterogeneity in gender gaps, including with the use of other machine learning techniques, but I believe this thesis may be the first to use the – increasingly popular – CRF technique.

The first part appears as a knowledgeable summary of the literature, but it is overall hard to understand if it does not merely paraphrase an econometrics/machine learning teaching book. Also the first part often assumes a quite advanced pre-existing knowledge of the reader about econometrics/machine learning techniques. Given that the techniques Aslan focuses on (Causal Random Forest) is becoming fairly standard in applied economics research, the author could have started his explanations at a more basic level, perhaps dropping instead his discussions of other ML topics that are only tangentially relevant to CRFs.

The second part is an interesting exploration of heterogeneous gender gaps using Causal Random Forests. One reason why this technique may have not been applied to gender gaps exploration, at least to my knowledge, is that it may not be obvious that gender gaps are akin to treatment effects. While the author acknowledges this, he could have delved deeper into the relevant discussion in the literature. Overall, this part feels somewhat like a demonstration of different CRF analysis tools on the question. However, it could have maybe been focused more on the analysis of a particular dimension of heterogeneity, and connect that result again back to economic theory that could explain why that dimension could matter.

Overall, this is a decent Master thesis, demonstrating that the student has a good understanding of what he does. However, there are also still quite a few typos, grammer mistakes, formatting issues, and also some errors in the thesis, suggesting a lack of effort in the final execution of the thesis.

CONTRIBUTION: 24 / 30
Application of a Standard Causal Random Forest Machine Learning mechanism to a published paper on gender gaps, to uncover potential heterogeneity in gender wage gaps. While other papers

have applied other machine learning techniques to estimate gender gaps, he seems to be among the first to use Causal Random Forests to test for heterogeneity in gender gaps.

METHODS: 27 / 30
Casual Standard Forest are relatively new, but increasingly popular, methods to empirical micro. In a way the whole thesis is a demonstration that the student knows and can apply this method. The thesis demonstrates that he indeed masters the method quite well.

LITERATURE: 20/20
Literature overview is complete and extensive.

MANUSCRIPT FORM: 14/20
Overall the thesis reveals that the students has not put in a large amount of effort to make sure it is error free and well formatted.

SUMMARY AND SUGGESTED QUESTIONS FOR THE DISCUSSION DURING THE DEFENSE:

Proposed questions for defense:
- Page 10, formula at bottom, last line: it's wrong? Should be $E[Y\_1|W=1] – E[Y\_0|W=1] + bias$?
- Pruning and penalization could be explained more. Why is it effective? How is pruning implemented? Section 2.2.2. is way too short. How is averaging done? Why does it introduce smoothness?
- Page 24: Are there downsides to using post-LASSO? Why would anyone estimate Lasso directly, or kitchen-sink OLS?
- Page 28: what are the weigths $g(x)$?
- Page 39: What is the "framework formulated by Huber and Solovyeva"?
- Page 44: is the "sequential conditional independence" assumption trivially fulfilled if the treatment is gender?
- How are the two variables in "heatmap" Figure 12 selected?

Please indicate whether you recommend the Thesis for defense or not.

I recommend the thesis for defense.

TEXT ORIGINALITY CONTROL

I confirm that I acquainted myself with the report on the originality of the text of the thesis from

[ ] Theses   [X] Turnitin   [ ] Ouriginal (Urkund)

Comments on the reported results: There were no suspicious aspects reported in the Turnitin report.

SUMMARY OF POINTS AWARDED *(for details, please see the page 3)*

CATEGORY		POINTS
Contribution	*(max. 30 points)*	24
Methods	*(max. 30 points)*	27
Literature	*(max. 20 points)*	20
Manuscript Form	*(max. 20 points)*	14
TOTAL POINTS *(max. 100 points)*		**85**
GRADE *(A – B – C – D – E – F)*		**B**

NAME OF THE REFEREE: **Andreas Menzel**

DATE OF EVALUATION: **23.08.2023**

_____

REFEREE SIGNATURE

EXPLANATION OF CATEGORIES AND SCALE:

CONTRIBUTION:
The author presents original ideas on the topic demonstrating critical thinking and ability to draw conclusions based on the knowledge of relevant theory and empirics. There is a distinct value added of the thesis.

Strong	Average	Weak
30	15	0

METHODS:

The tools used are relevant to the research question being investigated, and adequate to the author's level of studies. The thesis topic is comprehensively analyzed.

Strong	Average	Weak
30	15	0

LITERATURE REVIEW:
The thesis demonstrates author's full understanding and command of recent literature. The author quotes relevant literature in a proper way.

Strong	Average	Weak
20	10	0

MANUSCRIPT FORM:
The thesis is well structured. The student uses appropriate language and style, including academic format for graphs and tables. The text effectively refers to graphs and tables and disposes with a complete bibliography.

Strong	Average	Weak
20	10	0

OVERALL GRADING:

TOTAL	GRADE
91 – 100	A
81 – 90	B
71 – 80	C
61 – 70	D
51 – 60	E
0 – 50	F