Report on **Master Thesis**

Institute of Economic Studies, Faculty of Social Sciences, Charles University

Student:	Bc. Jan Provazník
Advisor:	RNDr. Michal Červinka, Ph.D.
Title of the thesis:	Consistency of S-weighted estimators in panel data models

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

Short summary

The thesis provides a detailed description of several robust regression estimators, such as Least Weighted Squares and the S-estimator, which are later combined into the S-weighted estimator. All estimators are compared on simulated data with different scenarios chosen for the exercise. The exercise compares different levels of data contamination for several models related to the S-weighted estimator.

Contribution

The contribution of the thesis is as stated in the proposal - it describes (history and theory) and compares robust methods for contaminated data. The execution of the simulation and comparison is clear and could be useful if someone needed to consult S-weighted and other robust estimators. It links the literature on contaminated data and explains strengths and weaknesses well. Nevertheless, the thesis is a simulation-based exercise.

The author contributes to the literature with a Python implementation of the algorithm. One might suggest putting it online on sites such as github.com so that others can benefit from it.

I find quite problematic that the title of the thesis states "in panel data models" and even though these methods as OLS and others can be used for panel data, I do not find a clear explanation when the methods presented and the simulation performed are useful for panel data problems and when not. If the thesis focuses on cross-sectional data, as I think it does, and the author's statement *"Note that this content differs from the original intentions described in the proposal. The complexity of the proof of S-weighted estimator's consistency in panel data models turned out to be an insurmountable obstacle while promising very little use, as the consistency in standard cross-sectional case has already been proved."* is true, why does the title still say "in panel data"? This distraction from the thesis and the problem is unfortunate and limits the contribution.

The study repeats Víšek's original texts in many places, and thus lacks further contributions.

Methods

The tools are well chosen for the purpose of the study, if we accept the exercise on non-panel data. Then, the author has well established the background and evaluated his hypotheses. As the thesis was written at an economic institute, one would suggest a sample of economic data for which the methods presented would be useful.

Literature

As above mentioned above, if we accept the non-panel data exercise, the author cites well the known literature and uses it to make arguments.

Report on **Master Thesis**

Institute of Economic Studies, Faculty of Social Sciences, Charles University

Student:	**Bc. Jan Provazník**
Advisor:	**RNDr. Michal Červinka, Ph.D.**
Title of the thesis:	**Consistency of S-weighted estimators in panel data models**

In the chapter 3, the author states towards the S-weighed estimation of Víšek (2015) that literature citing the work is limited and „To make things even worse, some of these articles are rather difficult to obtain." What does this mean for a reader? What is this literature?

Manuscript form

The manuscript is well structured and typed in LaTeX, which provides a good basis for readability. Unfortunately, there is a general comment on the reporting of tables and figures. In my opinion, the tables and figures in this thesis should be better described and labelled, as many have no label or description. However, the information is in the text, the flow of reading is slowed down by going back and forth to the text. For example, Figure 2.6. of Tukey's \rho function, it should have a legend or description with parameter values used to obtain the curves. Or, in Chapter 4 Simulation, most tables do not have a heading or description. For example, Table 4.6 shows the Kolmogorov-Smirnov test for what? There is no reference to this table in the text. In addition, the axes of graphs are not labelled with what we see in the figures.

Overall evaluation and suggested questions for the discussion during the defense

In my view, the thesis meets the requirements for a master thesis at IES, Faculty of Social Sciences, Charles University, I recommend it for the defense and suggest a grade D.

The results of the Turnitin analysis do not indicate (14%) significant text similarity with other available sources.

Q1: In the chapter 3, the author states towards the S-weighed estimation of Víšek (2015) that literature citing the work is limited and *"To make things even worse, some of these articles are rather difficult to obtain."* What is this literature, why is it so limited? What is the author's opinion on this?

Q2: How the consistency of S-weighted estimator is drawn from the simulation results?

Q3: Can you imagine an economic data set which analysis could benefit from the use of S-weigthed estimator?

Q4: Please see last paragraph in contribution section of the report.

SUMMARY OF POINTS AWARDED *(for details, see below):*

CATEGORY		POINTS
Contribution	*(max. 30 points)*	15
Methods	*(max. 30 points)*	22
Literature	*(max. 20 points)*	15
Manuscript Form	*(max. 20 points)*	10

Report on **Master Thesis**

Institute of Economic Studies, Faculty of Social Sciences, Charles University

Student:	Bc. Jan Provazník
Advisor:	RNDr. Michal Červinka, Ph.D.
Title of the thesis:	Consistency of S-weighted estimators in panel data models

TOTAL POINTS *(max. 100 points)*	**62**
GRADE **(A – B – C – D – E – F)**	**D**

NAME OF THE REFEREE: Luboš Hanus

DATE OF EVALUATION: 9. 9. 2024

Digitálně podepsáno (9. 9. 2024)
Luboš Hanus

_____
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.*

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.*

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

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.*

Overall grading:

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