

Report on Bachelor / Master Thesis

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

Student:	Bc. Jakub Černý
Advisor:	PhDr. Lenka Šťastná, Ph.D.
Title of the thesis:	Spatial Analysis of Czech Parliamentary Election: Comparison of Spatial Econometrics and Machine Learning

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

Short summary

The thesis aims to analyze determinants that explain the voting support of major political parties in the Czech Parliamentary election in 2021. The dataset includes voting results of 6,254 municipalities in the Czech Republic. Voting support of specific parties is often correlated in space, hence a spatial econometric approach is applied. The spatial model is compared to the machine learning approach, and the prediction power of both methods is analyzed.

Contribution

There is a vast amount of literature analyzing voting results in various countries. There are also some studies examining elections in the Czech Republic. Most of the papers, however, do not consider a spatial context that is highly relevant in this analysis. The author already focused on spatial analysis of voting support in his bachelor thesis (previous parliamentary election), but now applies a more rigorous spatial model to a greater set of control variables. Additional contribution is the application of the machine learning framework to this dataset and its comparison with the spatial econometric model. To my knowledge, this is the first analysis applying the ML approach to voting support considering the geographical context.

Methods

The author applies relevant methodological tools that are beyond the standard econometrics curriculum at the IES. He estimates a baseline model by the weighted least squares method. He then carefully tests for a specification of a spatial model (type of the model and also design of the weighting matrix). He runs the Spatial Durbin Error Model, which also controls for covariates of neighboring observations and spatial correlation of errors. Regarding the machine learning approach, the author starts with the Random Forest model in which spatial context is just incorporated by adding variables expressing geographical coordinates of municipalities and then continues with the Geographical Random Forest model, which is more computationally demanding but was explicitly developed for spatial data.

Literature

The author demonstrates his complete understanding and command of recent literature that focuses primarily on determinants of election results, spatial analyses, and novel machine learning approaches in the context of elections. The author quotes relevant literature appropriately.

Manuscript form

The thesis has a clear and logical structure. It is typeset in LaTeX. The student uses appropriate language and style, including the academic format for graphs and tables. The text effectively refers to graphs and tables and disposes of a complete bibliography.

Report on Bachelor / Master Thesis

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

Student:	Bc. Jakub Černý
Advisor:	PhDr. Lenka Šťastná, Ph.D.
Title of the thesis:	Spatial Analysis of Czech Parliamentary Election: Comparison of Spatial Econometrics and Machine Learning

Overall evaluation and suggested questions for the discussion during the defense

This thesis is an excellent piece of work. It was a great pleasure to supervise Jakub's thesis. Jakub carefully considered every step in the analysis and took a responsible approach to writing his thesis. In the thesis, he utilizes the knowledge he gained as a participant in the prg.AI Minor program.

In my view, the thesis fulfills the requirements for a master's thesis at IES, Faculty of Social Sciences, Charles University. I recommend it for the defense and suggest a grade A. The results of the Urkund analysis do not indicate significant text similarity with other available sources.

I have only one question for the defense.

1. You found that the prediction of the RF model is more efficient than the prediction of the GRF model. Do you have any explanation for that?

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

CATEGORY	POINTS
Contribution (max. 30 points)	28
Methods (max. 30 points)	30
Literature (max. 20 points)	20
Manuscript Form (max. 20 points)	20
TOTAL POINTS (max. 100 points)	98
GRADE (A – B – C – D – E – F)	A

NAME OF THE REFEREE: PhDr. Lenka Šťastná, Ph.D.

DATE OF EVALUATION: September 5, 2022

Digitálně podepsáno (5.9.2022)
Lenka Šťastná

Referee Signature

