



Diploma Thesis Evaluation Form

Author: Matěj Štěpánek

Title: Bc.

Programme/year: Bezpečnostní Studia (2.)

Author of Evaluation (supervisor/external assessor): Jan Kofroň

Criteria	Definition	Maximum	Points
Major Criteria			
	Research question, definition of objectives	10	10
	Theoretical/conceptual framework	30	30
	Methodology, analysis, argument	40	38
<i>Total</i>		80	78
Minor Criteria			
	Sources	10	10
	Style	5	4
	Formal requirements	5	5
<i>Total</i>		20	19
TOTAL		100	97



Evaluation

Major criteria: The master thesis deals with an interesting and policy-relevant topic – the precision of (European Values´) policy forecasting. While forecasting project of the European Values has been widely discussed (in general press as well as in journal “Obrana and strategie”), nobody has yet tried to evaluate accuracy (or precision) of this project. In this context authors analysis is original and much needed, especially if we consider that European Values wanted to use the project for formulating explicit policy-recommendations.

The paper has clearly defined research questions, and its structure is balanced and well-framed for the given purpose. The paper builds upon existing research dealing with assessing political predictions (mostly Tetlock and his various collaborators)

I do appreciate brief theoretical introduction into epistemology of predictions. Similarly, the paper benefits from standard and explicit methodology (using brier score as the key metrics). Also, I appreciate that the assessment has three parts – (i) internal assessment (analyzing precision of questions, logical consistency, consistency of invited forecasters vs. EV’s final predictions, etc.), (ii) comparison with other forecasting institutions (it is pity that only few questions could have been analyzed this way; nevertheless we can't blame the author here), and (iii) comparison against random guesses. Thus the analysis is robust and the conclusions seem valid.

From my point of view there are only minor points of criticism: 1) the author is right that the overall score of EV is too close to average random guess. Nevertheless it is still possible that the difference is statistically significant (a T-test could be informative here). While the EV’s Brier score is mediocre, it is still possible that the project has (even if very limited) added value.

2) There is a tricky point when the author compares frequency of assigned probabilities with the hypothetical random (uniform) distribution (see p 43, fig 1). If I were the author I would calculate expected values (frequencies for different probabilities) in a different way Author splits equally space between 0-1 (or 0- 100%) among eleven points – 0; 0,1; 0,2 0,9; 1 –thus creating equal intervals. My intuition, however, is that we need to realize that the extreme options “0” and “1” have only half of the surrounding intervals (0-0,05 a 0,95-1), as they are limited to their left and right respectively. Options like 0,5 or 0,7 have full surrounding intervals (0,45-0,55; 0,65-0,75). Thus expected values for two extreme options (0 and 1) should have been lower



than for the other nine options. Nevertheless it is highly likely that the main results of the comparison would have been almost the same...

Overall, results are persuasive and following recommendations for future forecasting projects stems directly from the analysis.

Minor criteria: The paper is logically structured around its research questions. Individual chapters, as well as its parts, are mostly well connected. Style is fine, even though the author is not a native speaker (here and there minor mistakes pop out, later chapters are a bit worse in this respect). It is certainly possible to argue that more economic style could save 10% of the text. Nevertheless, the argument is easy to follow, the author tries to be as explicit as possible and the text is rather well integrated.

There are numerous tables in the paper, these are easy to interpret (and the author always provides summarizing discussion in the text). A minor omission is that some tables miss upper-left cell (thus making a strange "dent").

It is great that the author provided well-structured appendix with all the data he used for his analysis.

Overall evaluation:

My impression is the thesis has been carefully prepared. It is analytically deep, original (at least in the Czech context) and relevant. Mistakes or errors are rare and substantially minor.

I am happy that the author was willing to engage with the mathematical aspect of predictions – and thus he was able to differentiate from the majority of our MA theses.

I dare to say the thesis belongs to top 5-10% of our MA theses.

Suggested grade: A

Signature:



**FACULTY
OF SOCIAL SCIENCES**
Charles University