

# Report on Master's Thesis

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

<b>Student:</b>	<b>Bc. Pavel Munduch</b>
<b>Advisor:</b>	<b>prof. PhDr. Ladislav Krištofek, Ph.D.</b>
<b>Title of the thesis:</b>	<b>Using spin model to determine FTTx connectivity market potential in the Czech Republic</b>

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

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

### **Contribution**

The thesis attempts to put together spin model which originated in Physics with analysis of motivation and potential of improvements of fiber cable technology infrastructure. This connection seems original and has not been discussed in the literature before.

Chapter 2 provides overview of Czech broadband market development till December 2019 (for some reason, period of 2020 is not included in the discussion). Section 3 presents two types of literature overview – selected research on broadband infrastructure and overview of applications of the Ising-type spin model in various fields of economics and finance. However, after the presentation of his own numerical results, the author does not compare his results nor provides an link to presented existing research, not even those results presented in Section 3.1, which is a bit of waste of potential. In Section 4, the author presents the methodology applied (see my comments in Methods section). In Section 5 and 6, the author presents and provides comments on his own numerical results. It appears that the R code and the four chosen initial state lattices are original contribution of the author. On the other hand, it is clear that the connection to the main economic topic of the thesis, in particular Czech market, is very vague, and in my opinion not properly justified, and the thesis seems to consist of three weakly connected parts (Ch3, Ch4 and the remaining chapters). Thus, the thesis seems contribute to analysis of dynamics of a particular form of Ising model rather than to provide answers to the three hypotheses formulated in the thesis proposal.

### **Methods**

The author uses Ising spin model, originally developed as a mathematical model of ferromagnetism in statistical mechanics. The basis of the model are discrete random variables with values +1 and -1 representing magnetic moments of magnetic spins, which are arranged in a lattice. The model considers interactions of each spin with its neighbors which may lead to a spin phase transition.

The presentation of Section 4 is corrupted with multiple misprints, mistakes and confusing/ambiguous notation (see my comments in Manuscript form section). It seems that some of the characteristics are of the author's invention, such as positive spin share, *dist* function, „risk cost eff“ and the spurious concept of „convergence“; at least the author does not provide any reference so these appear to be original. Those concepts and their presentation, however, strikingly point out to limited mathematical skills of the author, and, as such, are hard to evaluate as useful contributions. For some of the concepts, the author mentions alternatives, however, does not provide strong arguments for not implementing them; e.g. on the first line on page 30 he claims that *dist* function seems the most appropriate, but he does not mention based which criteria.

In the last part of 4.3.3 the author mentions a hypothesis test and formulates the null and alternative hypotheses. I did not find any further specification of the test procedure, nor could I find specific reference to outcomes of this test based on the numerical simulations.

# Report on Master's Thesis

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

<b>Student:</b>	<b>Bc. Pavel Munduch</b>
<b>Advisor:</b>	<b>prof. PhDr. Ladislav Křišťoufek, Ph.D.</b>
<b>Title of the thesis:</b>	<b>Using spin model to determine FTTx connectivity market potential in the Czech Republic</b>

To point out one essential shortfall, I am missing a discussion about the calibration of the model data (in particular parameters alpha and beta) based on the specific example of the broadband technology market in the Czech Republic as advertised. I would welcome at least a discussion about how to recover specific values of the parameters from the real data. As much as I am content with the interpretation of relative changes to the input parameters, the interpretation of specific values of parameters alpha and beta is not provided. The author claims to make such a discussion in Section 4.5 but I do not find satisfactory answers there.

## Literature

The author presents a decent literature overview in Chapter 3, although I do not find it appropriate to include specific results or their graphical representation of works where the authors applied the Ising model (or its modification) in a completely different research field. It has zero potential to be used as a benchmark for comparison with the author's own numerical results. As mentioned earlier, the author did not leverage from presented research results in Section 3.1 either, weakening the link between the content of Chapter 2 with Chapters 4-6..

The list of literature used in the manuscript suffers multiple shortcomings. Only about one half of the references seem to be peer reviewed publications. Majority of peer reviewed publications included in the list are presented with full information, except for Sznajd-Weron et al 2008 and Takaishi 2015 which do not include paging. Some of the remaining references state SSRN Electronic Journal, which is an online preprint database of working papers and abstracts before peer review publication. Thus, if the SSRN paper from 2011 is not yet published in a respectable journal with a thorough peer reviewing process, I would be cautious using the results of such manuscript. In some cases, the reference is so brief that it is hard to decipher the type of the publication, e.g. Dvorak 2012. I can only assume that some of the references without the website link are also online reports, e.g. UPC 2019. I would suggest to divide the list of references into two parts – journal papers, books or theses; and online articles or reports. The most spurious item in the list is LaMorte 2016 which is in fact a reference to an online teaching module on Probability theory [https://sphweb.bumc.bu.edu/otlt/MPH-Modules/BS/BS704\\_Probability/](https://sphweb.bumc.bu.edu/otlt/MPH-Modules/BS/BS704_Probability/) which I am safe to say presents (not only) the Central limit theorem with much less rigor than the bachelor's course Statistics JEB105 at IES.

## Manuscript form

As mentioned previously, the text of the thesis seems disconnected into three nearly standalone parts, Chapter 2, Chapter 3 and Chapters 4-6. Moreover, the readability of the manuscript is severely affected by multiple factors, see below.

The text is contaminated with frequent misprints, typos and grammar mistakes in English (in fact, also the text in Czech, cf. Abstrakt section), in some cases making it very challenging for the reader to decipher the content. The author even seems not to be consistent in use of definite and indefinite articles across the manuscript.

Mathematics presented in the manuscript is of really poor quality (from the notational and content point of view).

- Mathematical expressions are frequently displayed as a regular text, and even within the same sentence the author displays it in the font of regular text and slanted as well.
- There is an attempt at non-unique notation (e.g.  $J$  and  $J_m$  is presented as the same variables).
- Some expressions are corrupted to the degree one can hardly trace back the intended purpose or its correct form (e.g. the defining equation of the positive spin share on page 28 or ambiguous expression  $10^5/2/400$  on page 26). Further, the author presents lots of „copy-

# Report on Master's Thesis

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

<b>Student:</b>	<b>Bc. Pavel Munduch</b>
<b>Advisor:</b>	<b>prof. PhDr. Ladislav Křišťoufek, Ph.D.</b>
<b>Title of the thesis:</b>	<b>Using spin model to determine FTTx connectivity market potential in the Czech Republic</b>

paste“ typos in mathematical expressions, e.g. claiming  $s_i(t) = 1$  for positive spins as well as for negative spins in the first paragraph of 4.5; similarly for mathematical expression on page 61.

- The author abuses his own notation (e.g.  $t$  is used both in the sense of an index counting a time stamp of a potential spin change (I can only assume that it is running integers up to the chosen maximum time index) and at the same time  $t$  denotes the total number of potential spin changes; in definition of function  $dist$  indices  $j$  denotes position on the  $k \times k$  lattice while at the same time indices  $i$  denote neighbors of position  $j$ ; first the notation of the neighbors is confusing as e.g. for  $j=3$  one will end up comparing the same values  $s_3(t)$  of spin of a point at position 3 and  $s_3(t)$  of spin of its third neighbor; not to mention that it completely disregards the periodic boundary conditions).
- The author demonstrates poor knowledge of Central limit theorem and its assumptions (see the last paragraph of 4.2.1 on page 27 in which it is not clear, what objects form „our generated series“ and what are the properties of elements of such „series“ (I guess the author meant a sequence), neither it is clear what specific objects are the „statistics of our generated series“).
- The content of subsection Empirical approach of 4.3.2 is nonsensical – the author still did not explain what objects form „the series“  $a_n$ ; there is no interpretation provided for  $D(a_n)$  nor  $T$  and whether  $[1, T]$  is an interval which however makes no sense as it is claimed that  $[1, T]$  is a natural number. Expression (4.5) is not an equation as the author claims. The modification of definition of the limit of a sequence of real numbers hardly provides a limit of a sequence when he replaces „for every epsilon positive there exists natural index  $n_0$  such that for every index  $n$  greater or equal to  $n_0$  certain inequality holds“ by „for every epsilon positive and for every natural index  $n_0$  there exists index  $n$  LOWER or equal to  $n_0$  such that for every index larger or equal to  $n$  certain inequality holds“. The ill specification of the concept of limit can be illuminated in terms of the example linked to Figure 4.4 since selecting  $n_0=63096$  will result in a claim that the sequence diverges for the same selected value of epsilon while  $n_0=63097$  will generate  $n=63097$  and convergence. Further, as the author assumes only one fixed value of epsilon, he may get an affirmative answer even if the sequence does not stabilize but merely keeps on oscillating in the specified total range. Furthermore, I fail to see the link between  $n_0$  and the total potential spin flips.
- The author refers to several functions in the manuscript as metrics. This shows lack of knowledge of metrics in the mathematical sense as a mapping which has to satisfy certain axioms, in order to induce a topology on a set. One such false „metric“ is „risk cost eff“ defined in 4.3.4 for which the author does not give any motivation, interpretation, and not even a full title. To increase the level of confusion, the formula for „risk cost eff“ includes both value of  $Q(t)$  which I understand as a real value from  $[0, 1]$  and quantiles of  $Q(t)$ . I fail to see  $Q(t)$  as a random variable and even if I accept that it is a random variable, so must be the „risk cost eff“ as well.

The list of acronyms is vast and yet incomplete (e.g. CAGR, ESIF1).

Presented Table 4.1 contains a mistake violating the basic concept of probability, cf. the fifth case.

Presented Figures are mostly illuminating. E.g. Figure 2.1 and 2.2 are completely identical, although I do not believe it is by intention. In Figure 2.3 I would require more comments about its content, in particular what is the meaning of years below the percentage values in each column and why do they occasionally appear multiple times in different columns? In some cases, the legends included in the image is written in such a small script that it is impossible to decipher without a magnifying glass, e.g. Figure 5.5. Occasionally, the figures are confusing, e.g. left bottom image of Figure 5.8 should display

# Report on Master's Thesis

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

<b>Student:</b>	<b>Bc. Pavel Munduch</b>
<b>Advisor:</b>	<b>prof. PhDr. Ladislav Krištofek, Ph.D.</b>
<b>Title of the thesis:</b>	<b>Using spin model to determine FTTx connectivity market potential in the Czech Republic</b>

proportions but the range of values is from -1 to 0. Moreover, it is hard to compare the presented graphs when the colors used stand for different values in each image of Figures 5.1 to 5.20.

## Summary and suggested questions for the discussion during the defense

The results of the Urkund analysis **do not** indicate significant text similarity with other available sources. The Urkund analysis found just partial similarities to the earlier version of the author's manuscript.

In my view, the thesis fulfills the requirements for a master's thesis at IES, Faculty of Social Sciences, Charles University. Taking into account the above, I recommend the thesis for the defense and I suggest to award the thesis **grade E (51%)** depending on the performance during the thesis defense.

Suggested questions for the defense:

1. To what degree are the results sensitive to various values of the dimension of the lattice  $k$ . Did you run the experiments only for  $k = 1000$ ? Would  $k = 100$  or  $250$  be suitable as well?
2. You mention that „after consideration“ you decided not to consider the „diagonal neighbors“. What were the arguments for such a decision?
3. You consider binary variables  $J_i$  with values either 1 or 0. How would the model change if you change the binary values to  $\frac{1}{4}$  and 0, since you consider only four nonzero values for each position on the lattice and resulting in sum of  $J_i$  with respect to all positions in the lattice equal to 1?
4. What is the total theoretical range of values of parameters alpha and beta? Can alpha be negative, can it be arbitrarily large positive? You restrict yourself to interval  $[0, 10]$  based on inspiration in the literature, but what are the theoretical limits of the Ising model?
5. Why does Chapter 2 contain information only up to December 2019 included and does not extend to 2020? From this perspective the content of subsection „3600-3800 MHz auction (5G suitable frequencies)“ and some remaining parts of Section 2 appear outdated, not to mention that it ignores the recent auction of frequency domains (700 MHz and 3.4-3.6 GHz) announced in August 2020 and executed in November 2020.

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

CATEGORY	POINTS
Contribution (max. 30 points)	14
Methods (max. 30 points)	22
Literature (max. 20 points)	9
Manuscript Form (max. 20 points)	6
<b>TOTAL POINTS</b> (max. 100 points)	<b>51</b>
<b>GRADE</b> (A – B – C – D – E – F)	<b>E</b>

**NAME OF THE REFEREE: RNDr. Michal Červinka**

**DATE OF EVALUATION: January 22, 2021**

---

**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