Report on Master Thesis

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

Student:	**Vyacheslav Lypko**
Advisor:	**PhDr. Jozef Baruník, Ph.D.**
Title of the thesis:	**Modelling Dynamics of Correlations between Stock Markets with High-frequency Data**

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

The master's thesis "Modelling Dynamics of Correlations between Stock Markets with High-frequency Data" focuses on modeling realized correlations between selected stocks and commodities and forecasting accuracy of neural networks (NN) compared to more standard methods, namely autoregressive (AR) and heterogeneous autoregressive (HAR) models. This specific connection between neural networks and realized correlations is applied for the first time. Even though the author must have spent a lot of time dealing with high-frequency data and calculation of the realized measures, the thesis still looks as it were done very hastily. It is easy for the reader to get lost in the text as the author frequently jumps from one topic to another. Importantly, there are several statistical issues which raise serious questions about the results, such as the non-stationarity of the realized correlations. This problem is touched only slightly and it surely should have been given more attention. If the analyzed series are indeed non-stationary, then the comparison of neural networks with AR and HAR does not make sense because these are meant for stationary series. All in all, the depth and the focus on the empirical part should have been taken more seriously.

Several specific problems (not all of them are listed) in chronological order of the text flow:

- p. 5: It is really non-standard to have equations in the Introduction part.
- Sec. 2.3: DF test has a null of $d=1$ against $d<1$ which is not non-stationary vs. stationary series case, it is rather unit-root vs. no unit-root.
- Sec. 2.4: Why are only the AR models introduced, and not the MA models?
- Sec 2.4 is in general very badly structured and raises a question whether it even has its place in the text.
- p. 15: "microstructure noise…do not allows", English is sometimes very incorrect.
- p. 16: Author says that he doesn't expect NN to outperform HAR. Why? There is no reason given. Author has just probably stated so because the results have implied so. However, since the series of realized correlations were shown to be non-stationary, the comparison is quite questionable.
- Sec. 3.2: There is just a single reference in this section while the literature on the topic is very broad. Also, "integrated variance" is not defined anywhere.
- p. 21: The introduction to the chapter doesn't fit to the rest of the chapter.
- Are Figs. 4.1, 4.6 taken from somewhere? There is no reference and it is evidently copy/pasted (resolution of 4.6 is just awful).
- Tables 5.6, 5.7: Why no test for equality of correlations?
- Tables 5.9-5.14 should have been in the Appendix
- Why is R^2 from M-Z regression not reported? Some of the results are really non-standard, such as negative beta from M-Z regression.
- Summary trivializes the results.
- The thesis seems to be rather short. If the Figures and Tables were in the Appendix, I guess the length goes well below 60 pages for the diploma thesis.

For the reasons given above, I propose **B** as a final mark if the thesis is successfully defended.

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

CATEGORY		**POINTS**
Literature	*(max. 20 points)*	12
Methods	*(max. 30 points)*	23
Contribution	*(max. 30 points)*	18
Manuscript Form	*(max. 20 points)*	10
TOTAL POINTS	*(max. 100 points)*	**63**
GRADE	**(1 – 2 – 3 – 4)**	**2**

NAME OF THE REFEREE: PhDr. Ladislav Krištoufek
DATE OF EVALUATION: 6.9.2012

Ladislav Krištoufek

Referee Signature