

Report on Master Thesis

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

Student:	Bc. Sabyrzhan Tyuleubekov
Advisor:	doc. PhDr. Jozef Baruník, Ph.D.
Title of the thesis:	Can Model Combination Improve Volatility Forecasting?

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 and Methods

The thesis concerns various available possibilities for volatility (estimation and) forecasting. Various volatility proxies, forecast accuracy measure and training sample length is considered and illuminated for numerous models (both traditional including, among others, moving averages type, ARIMA type and GARCH type models, and several neural network type models) and several combination techniques.

In Section 2, the author presents the literature overview on volatility forecasting, existing results on use of modern machine learning techniques (namely variants of neural networks) in this field and on forecast combinations.

In Section 3, considered data set of stock daily data from yahoo Finance for 20 stocks (period January 2015 to January 2018) and four considered volatility proxies are presented.

The text of Section 4 on its own is quite a feat, as the author managed in a very fresh manner present a gargantuan overview of used methods. I particularly enjoyed this part for a clear mathematical description accompanied by essential and yet concise description of each particular model.

Section 5 containing the numerical results of the author is monumental. In total, the author compares the performance of 2400 forecasts generated on the considered dataset, subsequently aggregated and selected, ultimately resulting in overview of 48 sets of results. I welcomed the inclusion of Section 6 which serves as a summary of Section 5 as it was indeed a bit hard to get a global picture from the long (but necessary) overview of particular results in Section 5. Only then the answers to posed hypotheses in the introduction of the thesis became clear.

I am quite confident that the author will edit some parts of the thesis in a standalone paper which could be considered for publication in some international refereed journal on finance.

Literature

The author provides a large list of relevant literature, focusing mainly on peer reviewed literature sources (books or papers in respectable journals) and R package manuals. The list of used references follows a flawless citation style.

Manuscript form

The thesis is written in near flawless English, corrupted only with occasional typos which do not affect readability of the text. The manuscript is well structured and the author builds up the topic and provide smooth transitions between the main sections of the thesis so that the reader can proceed quickly through the main results. The text is well ballanced with reasonable number of tabullars and (illuminative) figures. Mathematical formulas are skillfully written in LaTeX which contributes positively to the superiority of this academic text.

Report on Master Thesis

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

Student:	Bc. Sabyrzhan Tyuleubekov
Advisor:	doc. PhDr. Jozef Baruník, Ph.D.
Title of the thesis:	Can Model Combination Improve Volatility Forecasting?

Summary and suggested questions for the discussion during the defense

To summarize, the manuscript accomplished goals formulated in the thesis proposal (although some of them have been replaced by different research questions) and the level of detail of the analysis of the topic makes it an interesting and illuminative text to read. As mentioned above, it even has a potential to be published in a refereed journal. As such, it is my honor to suggest **grade A**.

Suggested questions for the defense:

- 1) In the introduction part you state that the training sample needs to cover both periods of high and low volatility regimes for a stock. Is there a simple technique how to check that the selected training sub-sample satisfies such criterion?
- 2) Your data set is based on mainly the US stocks with some European stocks (e.g. MANU). Are the achieved results in the thesis sensitive to location of the market (e.g. considering US, European and Japanese stocks)?
- 3) Could you illuminate the reasoning behind the choice of parameters in High-low Parkinson proxy ($4 \ln 2$) and Garman-Klass proxy (0.5 and 0.39)?
- 4) How much computation time would have been required for the originally intended third training sub-sample of 2000+ observations? It is reasonable to pose a hypothesis that in such a case the KNN regression on combined fit would be even more dominating technique among all considered?

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

CATEGORY	POINTS
<i>Contribution</i> (max. 30 points)	30
<i>Methods</i> (max. 30 points)	30
<i>Literature</i> (max. 20 points)	20
<i>Manuscript Form</i> (max. 20 points)	19
TOTAL POINTS (max. 100 points)	99
GRADE (A – B – C – D – E – F)	A

NAME OF THE REFEREE: RNDr. Michal Červinka, Ph.D.

DATE OF EVALUATION: August 16, 2019

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.*

<i>Strong</i>	<i>Average</i>	<i>Weak</i>
30	15	0

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.*

<i>Strong</i>	<i>Average</i>	<i>Weak</i>
30	15	0

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

<i>Strong</i>	<i>Average</i>	<i>Weak</i>
20	10	0

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.*

<i>Strong</i>	<i>Average</i>	<i>Weak</i>
20	10	0

Overall grading:

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