Report on Master Thesis

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

Student:	**Bc. Jindřich Jurkovič**
Advisor:	**PhDr. Jozef Baruník, Ph.D.**
Title of the thesis:	**Forecasting Realized Volatility Using Neural Networks**

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

The submitted thesis "Forecasting Realized Volatility Using Neural Networks" deals with the state-of-the-art topic of realized volatility in connection with still rather unorthodox (at least from the finance mainstream point of view) neural networks. The main findings of the thesis are twofold – neural networks dominate the standard methods (on this point later) in-sample, and neural networks do not yield statistically significantly (based on Diebold-Mariano test) better predictions than the standard methods out-of-sample. These are evidently connected to the possible over-specification of the neural network methodology which yields a very good in-sample performance (the series are almost perfectly fitted due to high number of parameters) but only an average one in out-of-sample.

There are three major issues I have found. Firstly, the structure of the thesis is not satisfying. Separating the text into nine chapters/sections is not justified and it is quite disturbing for such a short text. It seems that the author either has not been able to structure the text sufficiently (and efficiently) or he tried to lengthen the text so that it meets the minimal page criteria of the Faculty. Connected to this issue, some sections are not necessary (Least Squares, Regression Analysis, even parts of Combined AR(I)MA Models) but some are insufficiently described (mainly the Neural Networks section but also the Volatility section). It needs to be stressed that OLS, regression and ARMA are standard methods so that they do not need to be described too much but Neural Networks are not standard so that these should have been given a proper attention and space. Secondly, neural networks are compared to "standard" forecasting techniques of volatility out of which the author uses ARMA and HAR (plus the HARD model is introduced). For no specific reason, the author omits ARFIMA which is a crucial mistake in the volatility forecasting as it is well known (and agreed on in the financial econometrics literature) that volatility is a long-term memory process (or a combination of both short- and long-term memory processes). Standardly, ARFIMA is the benchmark model. This is also connected to the missing literature review of volatility forecasting literature (had it been done, the author would have found ARFIMA being in practically all relevant papers). And thirdly, the (non-)stationarity-of-volatility issue is masked and bypassed by the author. In the methodological parts of the thesis, the author states that he needs stationary series for his analysis and then, in the empirical part, simply chooses to test stationarity on the first differences of the realized volatility. No reason is given. More problematically, the author then proceeds with forecasting the original realized volatility series. However, if the proper tests (ADF, PP and KPSS) tests were applied on the original series, the author would most likely find that the analyzed volatility series are neither unit-root (ADF and PP) nor stationary (KPSS) leading to series with fractional differencing parameter between 0.5 and 1. Even though this does not invalidate the forecasting exercise of the author, it should have been discussed and treated adequately.

Minor comments (chronologically with the text):

- p. 2: HAR is not LRD but it only resembles it.
- p. 4: "Volatility" is not variance but standard deviation.
- Sec. 2: Price process definition is missing.
- Sec. 2: Author keeps referring to "volatility of a price process" while it is "volatility of returns process".

Report on Master Thesis

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

Student:	**Bc. Jindřich Jurkovič**
Advisor:	**PhDr. Jozef Baruník, Ph.D.**
Title of the thesis:	**Forecasting Realized Volatility Using Neural Networks**

- Why is Eq. 2.3 used rather than Eq. 2.4? Why is Eq. 2.4 even mentioned when it is not discussed further?
- Sec. 4.1: It is not evident whether d is an integer or not.
- p. 25: Why is function labeled as L(x) in Eq. 5.3? In Eqs. 5.1 and 5.2, there are f_O and f_i.
- Sec. 5: How exactly are the weights of neurons estimated? What are possible input variables?
- Sec. 6: What is the analyzed period? It is not noted in the text, just the tables.
- Having sections for specific FX pairs is inappropriate.
- Is the KPSS test run on volatility or differenced volatility?
- How does stationarity of the differenced RV matter?
- p. 41: What is meant by "covariance horizon"?
- Sec. 7.2 of 5 lines?
- Sec. 8.1: How are the specific lags for MA extensions selected?
- Sec. 8.1: Notation of the specific models is very sloppy and confusing, and could have been done much more efficiently.

Overall, the thesis covers a very ambitious topic and probably brings interesting results. Unfortunately, these are hidden in rather chaotic and poorly structured text with a lot of unanswered questions (being it the previously mentioned ones plus the estimation of NNs or choice of MA parameters to name just two). **The thesis is defendable but I believe that it does not deserve a better grade than "Good" (Grade B). Therefore, unless the following questions are answered during defense, I propose "Satisfactory" (Grade C).**

1) **Why is ARFIMA not utilized in the text?**
2) **How are the reported specifications in Sec. 8 selected? What is the criterion?**

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

CATEGORY		**POINTS**
Literature	*(max. 20 points)*	12
Methods	*(max. 30 points)*	17
Contribution	*(max. 30 points)*	20
Manuscript Form	*(max. 20 points)*	9
TOTAL POINTS	*(max. 100 points)*	**58**
GRADE	**(1 – 2 – 3 – 4)**	**3**

NAME OF THE REFEREE: PhDr. Ladislav Krištoufek

DATE OF EVALUATION: 19.6.2013

Ladislav Kristoufek
cn=Ladislav Kristoufek,
o=Academy of Sciences of the Czech Republic, ou=Institute of Information Theory and Automation,
email=kristouf@utia.cas.cz,
c=CZ
2013.06.19 12:33:47 +02'00'

Referee Signature