Review of Doctoral Dissertation

Building and evaluating the Probabilistic Inflation Forecasting Expert System (PIFES)

Author: Daniel Robert Michael SONETT
Supervisor: dr. hab. T. Kamińska, prof. UG
Auxiliary supervisor: dr M. Brycz

I have been honoured by the Faculty of Economics, University of Gdansk to prepare the review of the PhD thesis. Below I present the detailed report on the quality of the research presented in this thesis, first assessing the general remarks and then presenting the detailed comments. I evaluate the academic quality and the contribution to the field, by answering three questions: (1) how important (and new) is the topic? (2) is the research and its description solid? (3) does the thesis provide enough contribution to the science?.

This study proposed the empirical tool, called PIFES, aimed for inflation forecasting. The methods used in the thesis come from the rapidly developing field of 'data science' (the intersection of econometrics and computer science). The thesis was prepared in English.

Topic and the structure of the thesis

The principal goal of the thesis was to build the expert system to forecast inflation. The system was built by applying machine learning methods and analysed a discrete variable defined by the range within the current inflation lies. I must admit this is a very demanding and ambitious task. Such study required deep knowledge of econometris and machine learning as well as some background in macroeconomics.

The thesis is composed of four chapters. The first chapter provides an overview of the applications of expert systems and maximum entropy methods. Chapter 2 discusses the foundations of PIFES, including the ones from the perspectives of machine learning, time series econometrics and theory of macroeconomics. Further chapter shows the details of the construction of PIFES (data sources, preprocessing, steps needed to prepare the forecast). Chapter 4 contains the evaluation of the forecasting performance of the system.

Literature used

The number of sources cited in the thesis are above the average standard of the PhD thesis. The choice of the literature is always subjective, as a result sometimes I missed some seminal papers, but my overal assesment of the quality of the literature is good.

The introductory parts present nice overview of the literature, e.g. on the idea of the entropy, its applications, stylised facts on the forecasting the inflation as well the variables typically used in inflation forecasting, among others. This extensive introduction makes the issue more readable for non-experts. However, sometimes I disagree with the specific points. More specifically, the thesis did not notice numerous application of entropy in finance (as pioneered by Hansen, Sargent 2011). Feature selection methods have been already used in inflation forecasting (e.g. Bai, Ng 2008; for Polish inflation: Szafrański 2013); this thesis does not consider this strand of literature, although applications for differend fields of science are mentioned; even more, the author states: "So far, the explanatory variables within these [inflation forecasting] models, have not been examined."

(page 75). Generally, the author tends to use interdiscipliary literature and sometimes I had a feeling that the author could have had a closer look at the literature on economics, instead of differend fields.

To sum up this point, the thesis makes good use of the literature.

Formal issues

The thesis is clear, in a way that after reading the thesis I was able to understand and evaluate the whole research process. However, I found a numerous misprints (e.g. 'neuron' rather than 'neutron', page 83; 'integation' rahter than 'integrity', page 125), ocasionally even inside the mathematical formulas (e.g. all formulas 0.2-0.4 contain x_1 while it should be: x_1, x_2 and x_3 respectively; similar typo occurrs on page 133, formulas 1-5, where *range1* should not appear in 2-5). Morevoer, sometimes the word order in the sentence was confused.

The style of literature references is not consistent—as a representative example I may point out page 77 of the thesis, where threee styles of referencing appears: second name only (Chang 1997), initial + second name (J. Binner et al. 2005) and second name only in capital letters (ARTIS et al. 1995).

Generally, the formal issues are the weakest point of this thesis.

Research methods

Data mining model basically are aimed to model (and/or forecast) the continuos variable (regression based methods) or discrete variable (classification). This thesis takes the second perspective and uses five cattegories. While classification methods are typically used in microeconometris and rarely in macroeconometrics (especially when analysing time series for a single country or region). The author provided a motivation for using classification rather than regression based method (e.g. on page 161). However, I missed more emprical comparisons between these two approaches.

The empirical part of the thesis mostly relied on testing the out-of-sample forecasting performance of the PIFES. Empirical data (annual time series) were collected from one source – St. Louis Fed. I have no concerns with respect to data sources or preprocessing. Verification of PIFES was performed sufficiently. Two benchmark models are choiced: artificial neural network ('machine learning' method, henceforth: ANN) and OLS (traditional econometric method). The forecast evaluations relied on two measures: accuaracy and RMSE. While the first is typically used in classification, the second measure is not fully relevant for classification (even if the variable is ordered, as in the case of PIFES). In my opinion, providing a classification table could provide more detailed, sythetic information on the quality of forecasts (one could, however, take a closer look on the mis-classifications by analysing e.g. Figure 4.9).

To sum up, even though I have some remarks to the methods, the methods were chosen properly and the overall quality of the research is fair.

Findings and value added

The PIFES has been found to be comparable, in terms of forecasting quality, to the OLS and ANN. I think one should focus on 1-year or 2-year ahead forecast, while the author considers also long term forecasts. In the case of short term forecasts, PIFES slightly overperforms benchmarks, which confirms the relevance of PIFES. Surprisingly, this empirical finding was not highlighted.

My strong concern is very brief description of the benchmark models. The benchmark models used similar starting points and they are based on the same dataset. However, I have not found an exact model specification (for OLS) or the details on the structure of the ANN. For instance, OLS could mean this is AR(1) model or the model using some (which?) of the independent variables used for PIFES.

When looking at the results, one can notice that for 2013-2016 both OLS and PIFES provided perfect forecast accuracy (see Table 4.1, lowest panel). I regret this finding and its practical consequences have not been further investigated.

In addition, I missed comparison of the forecast from PIFES to the 'naive' (random walk) benchmark, that is assuming that the 'range' in *t+1* is the same as it was in *t*. Such a benchmark is typically used when assessing out-of-sample properties of time-series or cross-sectional time series (e.g., the paper by Stock and Watson 2008, cited in the thesis). I wish we could discuss this additional issue (or even discuss some new findings) during the thesis public defence.

I agree that PIFES provides accurate forecasts and may be directly applied for inflation forecasting. However, when using annual series one may use for forecasting the year *t+1* (say, 2019) only when all the source data on the year *t* are available. This condition is typically fulfilled in April or May. The data availability provides a natural limit for real-life applications, in the case of PIFES the forecast may be prepared no sooner than in April/May. Unfortunately, this issue was not discussed in the thesis, although the extensive discussion is provided on the possibility to extend PIFES.

Chapter 4.2 lists the possible extensions of PIFES. I agree with these possibilities. However, I expected a more detailed discussion of the results and providing the reader with the possible reasons for such results.

To sum up, the methods used in the thesis are routinely employed in 'data-mining' microeconometric application, but the application of the expert system for inflation forecasting is very original and ambitious task. In my opinion, by building and evaluating PIFES, Mr. Sonett provided a valuable contribution to the field of economics.

Detailed comments

Below I present detailed, usually minor comments.

- Kullback (as in the thesis) or Kullback-Leibler divergence (as mostly cited in the literature)?

- GMM method is asymptotically efficient (see Hall 2005, pages 88-90; also this fact was mentioned on page 45 of the thesis) and in some cases it is equivalent to the maximum likelihood (e.g. under standard OLS assumptions + normality of error term); in this case I do not understand questioning the properties of the GMM (page 45, first full sentence),

- formulas (2.12) and (2.13): $(X_i|Y)$ formally does not mean anything, so I would prefer $P(X_i|Y)$ notation,

- in my opinion there is not need to providkażde badanie ma swoje wadye a formal definitions of the concept well established in the field of applied economics. e.g. time series or lag operator (Def. 2.2.1.1 and 2.2.3.1., page 100 and 104, respectively),

- OLS does not assume normal distribution of the error term (see, e.g., Kennedy 2003, chapter 3.2), at least for point estimates.

- stationarity: the conditions of weak stationarity of AR(1) process (p. 103) hold only for $a>0$, which is not stated in chapter 2.2.2.

Conlusions

In this report, I focus on three major criteria: (1) importance and originality of the topic? (2) quality of research and its description, and (3) the contribution to science. Below I sum up the assessment of these criteria.

Inflation forecasting is an issue relevant both from practical (e.g. central bank) and research perspective. Also, applications of machine learning methods become an increasingly popular research area. Therefore, I find the topic undertaken in the study as important and actual.

The research methods are sound, but already estaiblished in the literature. While the thesis does not propose a new methodology, the application of expert system to provide categorical forecasts of inflation is new in the literature. Hence, in my opinion, the research quality is fair and the doctoral thesis has an important contribution to science.

Obviously, every research has some limitations or drawbacks. Let me focus on two major comments. The author makes good use of the literature, however, the literature survey in not enough focused on the papers specifically related to economic. I also have some concern with respect to the benchmark models (very brief description, missing some simple methods like naive forecast). Nevertheless, my overall assessment of the thesis is high.

Based on this, I conclude that **the thesis fulfils all the legal requirements as well as the academic standards**. Thus, I recommend the Faculty of Economics approving for public defence.

Literature used in the report

Bai, J., & Ng, S. (2008). Forecasting economic time series using targeted predictors. *Journal of Econometrics, 146*(2), 304-317.

Hall, A. R. (2005). *Generalized method of moments.* Oxford university press.

Hansen, L., & Sargent, T. J. (2001). Robust control and model uncertainty. *American Economic Review, 91*(2), 60-66.

Kennedy, P. (2003). *A guide to econometrics.* MIT press.

Szafrański, G. (2013). Syntetyczne wskaźniki wyprzedzające w prognozowaniu inflacji w Polsce. *Przegląd Statystyczny, 60*(3), 395-416.

dr hab. Paweł Baranowski, prof. UŁ

04/04/2019