McGill
Department of Economics
McGill University
Leacock Building
855 Sherbrooke Street West
Montreal, Quebec, Canada H3A 2T7

Département de sciences économiques
Université McGill
Pavillon Leacock
855, rue Sherbrooke ouest
Montréal (Québec) Canada H3A 2T7

Telephone: (514) 398-3030
http://www.mcgill.ca/economics/

30 April 2024

External evaluator's report on the dissertation:

Essays in Time Series Forecasting

by Filip Stanek

This thesis presents highly interesting and technically accomplished work, and is very well written and presented. It has been a pleasure to read.

Two of the essays have already been published, and all three are well developed, so my comments will primarily concern possible directions for future research.

Chapter 1.

This chapter addresses the important case in which researchers divide a sample into training and test (pseudo-out-of-sample) subsamples. It is also common, however, to divide samples into three, called for example training, validation and test; the validation set is commonly used to optimize the value of a tuning parameter, such as the penalty parameter in the LASSO. In this way the researcher can use the test sample to evaluate a method which has been optimized in sample with respect to tuning parameters. This might be a natural extension for future research.

With respect to the empirical evaluation of performance: the author points out that some of the series used are unlikely to be stationary, which conflicts with the assumptions used in this chapter. However there is also the suggestion that this in some sense biases the evaluation against the author's method (this is an `adverse' setting), so that if only stationary series were used, its relative performance would tend to improve. I don't see however why this is the case. What do we know about the relative impact of departure from the stationarity assumption on different methods that are compared? I agree that there is some empirical evidence on the point is available from comparison of results on different series, some of which can be identified as likely to be locally stationary, but nonetheless it should be possible to make some theoretical exploration of the nature of the distortion when non-stationary (esp. I(1)) series arise.

How are the variance-minimizing weights obtained in practice?

Finally, another paper of Inoue and Kilian which it might be nice to mention in this context is `In-sample or out-of-sample tests of predictability: which one should we use?' (Econometric Reviews 2007); the relationship with the present research is perhaps a little distant, but it does make an interestingly counter-suggestive argument about in-sample evaluation.

Chapter 2.

This chapter is clear about the fact that marginal improvements are modest, which is commendable. Perhaps in future work some other field of application of regularization for methods such as these might yield greater improvements. I wonder in particular if this idea has been fully exploited in feasible GLS estimation? (By analogy to this research, OLS with equal weights represents one polar case, weighted least squares with diagonal elements only an intermediate case, and on to more general FGLS models of the covariance matrix.) FGLS has been somewhat out of fashion in recent years—researchers have generally fallen back on replacing least squares standard errors with HAC standard errors, abandoning the idea of improving the parameter estimates. But there has been some recent work trying to return to the idea of using FGLS for better parameter estimates; see for example Chaudhuri and Renault 2023:
https://warwick.ac.uk/fac/soc/economics/research/workingpapers/2023/twerp_1473_-renault.pdf

In any event it looks overall as though most of the information content for forecasting has been squeezed from past volatility. Is there potential to use regularization methods when combining past volatility information with additional variables that can be used to forecast?

A minor comment: I'm not sure I agree that the LASSO would lead to unmotivated [footnote 2] equality or otherwise of some elements; the motivation is the usual risk/bias tradeoff embodied in the LASSO. With high-risk, low-value additional parameters, the probability of getting a better estimate by including the variable, rather than leaving it out, can be well below 50%, at a given sample size. As with ridge estimation, assigning a too-large coefficient may increase risk more than shrinkage/elimination bias.

Chapter 3.

As in the previous chapter, this chapter addresses a case of using a smooth transition between polar cases with the aim of finding a point which produces better results than either pole.

This is interesting and clearly very successful work, but just a comment on the competition and incentives.

This situation is a little reminiscent of stock picking contest, in which for example a large group of students manages a notional portfolio and the top three get job offers from an investment firm. Clearly the incentive there is not to provide prudent investment management, but instead to take big risks, because the people who win will have to be people who both took big risks and were lucky. Good prudent management will put you upper mid- pack, which is of no use from the point of view of winning a prize.

The investment challenge part of the forecasting competition has some aspect of this feature. The author points out on page 89 that risk was regulated to improve the chances of securing a high rank. This is clearly the correct strategy, but it seems to me that its success is as much a comment on the author's general cleverness than on the performance of the model! Such artificial challenges do not correspond well with, for example, the criteria for model evaluation set out in section 3.2. The sinusoidal problem evaluation corresponds better, but it is again a somewhat artificial case from the point of view of economic forecasting.

This is of course a technical quibble in some sense. It's clear from the results at the end of section 3.5 that improvements are quite general, and that this model class is potentially a successful approach to a wide variety of problems.

To summarize, this is an excellent thesis, presented with the honesty in stating limitations that is characteristic of fine scholarly work.

In closing, let me state explicitly that the thesis clearly satisfies formal and content requirements for a Ph.D. thesis in Economics, and that I certainly recommend that the dissertation should proceed to the defence.

Sincerely,

John Galbraith
Professor of Economics
McGill University, Montreal
john.galbraith@mcgill.ca

Minor typos.

Chapter 1

--p.13, line after (1.23): generalilzed --> generalized

Chapter 2

--p. 67, line before (2.7): transforemed→ transformed
--also p. 67, last line `with the only the test' → `with only the test' (?)
--p. 68, three lines before Table 2.4: `This also likely explain'→`This also likely explains'