

Economics 4151

**Take Home Exam**

Due by Email on Wednesday, May 12 at 5:30pm

For this take-home exam, you will write a report summarizing your analysis of some time series data. You will make use of the techniques considered in the homework assignments. While the take-home exam is a bit like a homework assignment, you must work independently. In any event, everyone has different data, so your findings will not be the same.

**1. Report**

You have just been hired by a local manufacturer that, coincidentally, has the same name as your name (e.g., if your last name is “Smith”, then the firm is “Smith Industries”). You have no idea what they manufacture, but they didn’t hire you to work on the assembly line, so it doesn’t really matter (but don’t worry, I’ve been told that it is an ethical business). The firm hired you for your exceptional skills in applied time series econometrics. They heard from me about how well you’ve done on the homework assignments and they were looking for someone to help them forecast future sales.

The firm has lots of data on sales, compiled at a monthly frequency. They also have data on regional industrial production, the regional CPI, and sales for some other manufacturing firms in the area. Alas, no one at the firm knows the first thing about time series econometrics. Your task then is to write a report for your new boss explaining what forecasting model they should use for sales. The letter from the boss is given below:

April 30, 2010

Dear New Employee,

We are delighted that you have joined \*\*\*\* Industries. We are proud of our ethical manufacturing practices. We hope to stay ahead of our (less ethical) competition by forecasting our future sales better. Right now, we just use a rule-of-thumb that future sales will continue to grow at the same average rate they have grown in the past.

We heard from Professor Morley that you have exceptional forecasting skills. We would like you to apply your skills to figuring out the best approach to forecasting future sales. We then want you to prepare a short report (about five pages) on why we should follow your proposed approach instead of our current rule-of-thumb approach.

Here are some issues that we would like you to address:

- i. We are interested in forecasts at different horizons. We care about one month ahead. But we also care about the level of sales one year ahead. Again, we currently just use average past sales growth to forecast both. Please convince us that we should use something different.
- ii. We care about our forecasts being off in either the positive or negative direction equally. We also fear being way off much more than being close. While we don't know anything about time series, Professor Morley assures us that this means we have a "quadratic loss function". Please take that into account when determining the best forecasting approach.
- iii. One of our current employees is convinced that interest rates have a big impact on our future sales, although our ethical product only costs \$2.99 per item. We have included some interest rate data courtesy of Professor Morley. He says it is already in "real" terms, so you don't need to account for expected inflation. Please let us know if we should consider interest rates in our forecast.
- iv. We would like you to consider classical hypothesis testing as means of determining an appropriate forecasting model (at least that's what Professor Morley tells us to ask you to do).
- v. There have been a number of changes in local sales taxes over the last 45 years. Some local macroeconomists told us that we should worry about something called the Lucas critique and avoid forecasting based on "reduced-form relationships in the aggregated data". They suggested that we might survey our customers to get their expenditures on our product, their personal incomes, and any other data that might be relevant. Then we are supposed to use this information to develop a "structural model" of our sales based on "microfoundations and deep structural parameters". The cost of this data collection would be quite high and we decided to hire you instead. Do you think this data collection would be worthwhile? Explain your reasoning. Is there anything you can do with the current data to help answer this question?

We look forward to seeing your report by 5:00pm on Monday, May 5. In keeping with our ethical mission, you will receive a brief vacation after the report is due.

Sincerely,

The Big Boss of \*\*\*\* Industries

P.S. Please keep the report to about 5 pages. Avoid using too much technical jargon, except when necessary. Remember that we aren't familiar with time series econometrics. Pictures and (neat) tables are helpful, but they must fit in the report. No appendices please. Please submit by email as a pdf.

So that's the letter and your task. You can obtain the firm's data file from the class website. I've had a look at the letter and data and here are some thoughts/suggestions that you should consider in preparing your report:

- The sales data are in nominal terms. The strongest forecasting relationship may be in terms of real sales. Also, because the firm wants to know about future sales in order to set levels of production, they really should be interested in future real sales. You can use the regional CPI to deflate the nominal sales figures. (The industrial production series is already in real terms.)
- You should consider whether data series should be transformed into natural logarithms and whether they have unit roots and/or structural breaks. If so, this should all be accounted for when developing your forecasting approach. In particular, unit roots mean that you should difference the data. A structural break in, say, the mean of a series should be accounted for in your forecasting model using a dummy variable. Don't forget to tell the firm how to get back to a forecast of the level of real sales.
- To be convincing when comparing your approach to the firm's rule-of-thumb approach, it is a good idea to set aside the last 100 observations to be a "hold-out" sample. That is, do all of your model selection and parameter estimation for the first 400 observations. Then compare how well your proposed model does relative to the rule-of-thumb approach in terms of forecasting the last 100 observations. In particular, consider the root mean-squared-error (RMSE) of the forecast errors for different models over the last 100 observations. For simplicity, you can use real sales (rather than nominal sales) when considering the firm's rule-of-thumb approach. That is, make the benchmark forecast of future real sales growth equal to average historical real sales growth. Report relative RMSEs vs. the benchmark model.
- Be clear about what are all of the possible models that you considered and why you argue for one model over all of the others. If you get any ambiguous results using classical hypothesis tests, use the holdout sample to compare the different models and propose a model based on which does best in the hold out sample.
- When considering the possibility of spurious regressions, it can be useful to conduct bootstrap tests for significance. I.e., consider a model of sales without the potentially spurious variable for the BDGP. Then consider regressions using the simulated data and the potentially spurious variable to determine the distribution of the test statistic on the potentially spurious variable. Also, spurious regressors should not forecast well in the holdout sample. If you do not include variables such as the interest rate in your proposed forecasting model, explain to the firm why not. Note that you do not have to consider sales data for other firms.
- On the Lucas critique issue, look for evidence of structural breaks in slope coefficient parameters in the forecasting models. The Lucas critique is not simply about whether there is a structural break in any parameter, but specifically that the correlation between variables is unstable.

The organization of your results should be roughly as follows:

1. Transformation of data into a form that is suitable for modeling with stationary time-series models.

2. Estimation and testing for a set of possible models, including models with the interest rate as a predictive variable.
3. Out-of-sample RMSE results for different models to help settle on final proposed model (is it different than what was suggested by hypothesis testing). Provide some explanation of the relative RMSE calculations.
4. A quick summary of the proposed model and how to reconstruct a forecast of the level of real sales using it. Report confidence intervals for the model parameters.
5. A quick discussion of whether the Lucas critique is relevant using tests for structural breaks in forecasting model parameters (especially slope coefficients).

That's all of my advice for now. Good luck!