Instructor:
Dawit Zerom

CALIFORNIA STATE UNIVERSITY
FULLERTON

Mihaylo College of Business and Economics
Department of Information Systems & Decision Sciences

PROJECT #4

ISDS 526: *Forecasting for Analytical Decision Making*

For Students: Do not reproduce or transmit. I spend a lot of time and effort in creating useful materials for my students. You will benefit nothing by making this material available online or transferring it to other people. On the negative side, my effort becomes worthless and the material will no longer have any educational value.

Case: The American Automobile Association

AAA Washington is one of the two regional automobile clubs affiliated with the American Automobile Association (AAA) operating in Washington State. AAA Washington was founded in 1904. Its service territory comprises the 26 Washington counties west of the Columbia River. The club offers its members a variety of automobile and automobile-related services. Member benefits include emergency road services; a rating service for lodging, restaurants, and automobile repair shops; tour guides to AAA-approved lodging, restaurants, camping, and points of interest; and advocacy for legislation and public spending in the best interests of the motoring public. In addition to these services, AAA Washington offers its members expanded protection plans for emergency road service; financial services, including affinity credit cards, personal lines of credit, checking and saving accounts, time deposits, and no-fee American Express traveler's checks; access to a fleet of mobile diagnostic vans for determining the "health" of a member's vehicle; a travel agency; and an insurance agency.

Instructor:
Dawit Zerom

Issue

Club research has consistently shown that the emergency road service benefit is the primary reason people join AAA. Providing emergency road service is the club's single largest operating expense. It is projected that delivering emergency road service will cost 9.5 million US dollars in the next fiscal year, 37% of the club's annual operating budget.

Of particular concern for the club was a rise in the use of emergency road service by members. Membership had been growing steadily for several years, but the increased cost was more than what could be attributed to simple membership growth. Mr. DeCoria, the club's vice president of operations, then checked to see if there was a growth in emergency road service use on a per membership basis. He discovered that between fiscal year 1990 and fiscal year 1991 the average number of emergency road service calls per member grew by 3.28%, from an average of 0.61 calls per member to 0.63 calls (AAA fiscal year begins July 1). Concerned that a continuation of this trend will have a negative impact on the club financially, Mr. DeCoria decided to use his knowledge of business forecasting (he took an equivalent course to ISDS 526 when he was a student) to analyze the emergency road service calls volume and hopefully come up with an objective assessment of the situation.

Potential predictors or explanatory variables

Previous research done by the club discovered several factors that have an impact on emergency road service call volume (*Calls*). Among these factors are average daily temperature (*Temperature*) and the amount of rainfall (*Rainfall*) received in a day. This research has shown that emergency road service calls increase as rainfall increases and as average daily temperature falls. Mr. DeCoria has also observed that the cyclical trend of the time series seems to be lagging behind the general economic cycle. He has suggested that the unemployment rate (*Rate*) for Washington State would be a good surrogate measurement for the general state of Washington's economy. Data on all these variables is available.

Summarizing:

Calls: (Dependent/forecast variable)

Potential predictors:

Rate
Temperature
Rainfall

Your assignment

The role of your group, the forecasting consultants hired by the American Automobile Association (hereafter referred to as AAA), is to come up with a healthy, easy-to-use, accurate, and powerful dynamic regression model. The purpose of the model is to explain the variation within the provided service calls volume data (referred to as Calls in the model); hence allowing Mr. DeCoria insights into why service calls have surged at particular times. These insights will aid in Mr. DeCoria's decision making process and allow him to take appropriate measures to counteract the additional expenses of increased roadside service calls. To arrive at the final dynamic regression model, you will undertake several model building steps.

Organization

You will organize your report per the following **headings**, in that order. Of course the contents are for you to fill in. **THESE ARE NOT QUESTIONS** that you need to answer. THESE are checklist of things that your report should address.

1. Executive summary

2. The forecasting problem

3. Dynamic regression model building

Step1. Starting out

- Because the weather condition is a major determinant of road-side service calls, you first investigate how temperature and rainfall affect calls.
- Plot calls and rainfall and see how they behave over time. What type of relationship do you observe? Do the same for calls and temperature.
- Based on your observation from the graph, run two simple regressions where the first is calls on rainfall and the second is calls on temperature.
- Assess the two models based on their predictive power – use R-square.
- Now plot rainfall and temperature together. You will see that their pattern is quite correlated. What does this mean to regression?
- To answer this, run a regression of calls on both rainfall and temperature. Assess this model using R-squared, Adjusted R-squared and the significance of the coefficients.
- The above regression implies that perhaps rainfall does not add much to predictability of calls once we have temperature in the model.
- Label the model of calls on temperature – **MODEL #1.**

Step2. Capturing further seasonality

- As you can expect *calls* are seasonal by nature. MODEL #1 will capture those seasonal patterns due to weather conditions. *But what about those due to people's behavior such as long vacations in the summer*.
- To see whether or not MODEL #1 failed to capture all the seasonality in calls, you investigate the *error autocorrelations* of MODEL #1. Do you see spikes at multiples of 12? If yes, then MODEL#1 should be extended to capture this seasonality.
- To do so, we add seasonal dummies. When you add dummies you first need to decide the reference month. If you look into the plot of calls, you will notice that the month of September often records the lowest number of calls. To facilitate interpretation, use September as reference month.
- After adding the 11 seasonal dummies, assess the resulting model. As one might expect, not all the dummies will be significant because weather conditions have already accounted for much of the seasonality. Remove those dummy variables that are highly insignificant. Use the 10% level. Label the model of calls on temperature and significant seasonal dummies – **MODEL #2**

Step3. Examining and dealing with autocorrelation of errors

- To see whether or not MODEL #2 is affected by series correlation problem, examine its error autocorrelation. What do you observe? Do you see autocorrelation?
- Another way to see this is by using the Ljung-Box Statistics. If the statistics shows significance, then it means that autocorrelation may be seriously affecting the model.
- To address this problem and based on the pattern of the error ACF, one solution is to add a few lagged terms of the error – are called AUTO[-n] in forecast pro. This is the Cochrane-Orcutt solution we covered in class. Start by adding AUTO[-1]. What happens to the model? Is Ljung-Box statistics still significant? Also look at the pattern of the error ACF. Do you need more lags? If not, stop.
- Label your new model – **MODEL #3**.
- Now compare MODEL #3 with MODEL #2 by assessing several aspects of the models such as Adjusted R-squared, BIC, Ljung-Box statistics, MAPE, etc. Which model is preferred?

Step4. Further improvement

- I am sure you recall this statement "*Mr. DeCoria has also observed that the cyclical trend of the time series seems to be lagging behind the general economic cycle. He has suggested that the unemployment rate (Rate) for Washington State would be a good surrogate measurement for the general state of Washington's economy.*"
- To test Mr. DeCoria's argument, you decide to add the predictor **Rate** to MODEL #3. **One at a time, add <u>Rate itself</u> or <u>lags of Rate</u>. Do not add more than one Rate variable (lagged or otherwise) at a time as they are highly**

correlated with each other. Among all the options, you have tried, you will see that Rate[-5] is the best.

- Now what happens to the new model? You may notice that AUTO[-1] becomes less significant, but will still be significant at 10% level.
- Label the new model with Rate[-5] – **MODEL #4.** Now compare MODEL #4 with MODEL #3 by assessing several aspects of the models such as Adjusted R-squared, BIC, Ljung-Box statistics, MAPE, etc. Which model is preferred?

4. The final dynamic regression model

- Write the final model – MODEL#4 – in an equation format.
- Interpret all the coefficient estimates. Do they make business sense? Elaborate.

5. Forecast accuracy

- In order to assess the **genuine forecast accuracy** of your dynamic regression model, one has to test the model **using outside data not used in the model building process**.
- Luckily, we have calls, Rate and Temperature data for the next 5 months.

	Calls	**Rate**	**Temperature**
Sep 1992	20251.00	6.78	58.40
Oct 1992	22069.00	7.59	53.20
Nov 1992	23268.00	7.99	44.80
Dec 1992	26039.00	8.25	37.80
Jan 1993	26127.00	9.53	34.90

- To implement out-of-sample forecasting in Forecast Pro, **simply append the above data to the original data without the date column**. Then, open the data in Forecast Pro and apply your saved script for the specification of the predictors. Also make sure to holdout 5 observations and forecast 5.
- Provide a table showing the accuracy of your forecast using MAPE
- Compare the regression model with “time series only models”. To do so, provide MAPE for the “Expert Selection method” of Forecast Pro.
- Which method is more accurate? Your regression or Expert Selection?

6. Conclusion