

Use SAS/ETS to Forecast Credit Loss for CCAR

Xinpeng “Tom” Wang, Huntington National Bank, Columbus, OH

ABSTRACT

After the 2007-2009 financial crises, the Federal Reserve begins to conduct annual stress tests of Bank Holding Companies (“BHCs”) with total consolidated assets of \$50 billion or more (“Covered Company”). To estimate credit losses for BHCs loan portfolio, one solution is to employ SAS logistic and regression procedures to predict probability of default (PD) and loss given default (LGD). This paper presents an alternative utilizing SAS/ETS package to forecast net charge offs (NCOs) on the aggregated data. Then the paper discusses general limitations of models in financial risk management and specific shortages for both approaches.

INTRODUCTION

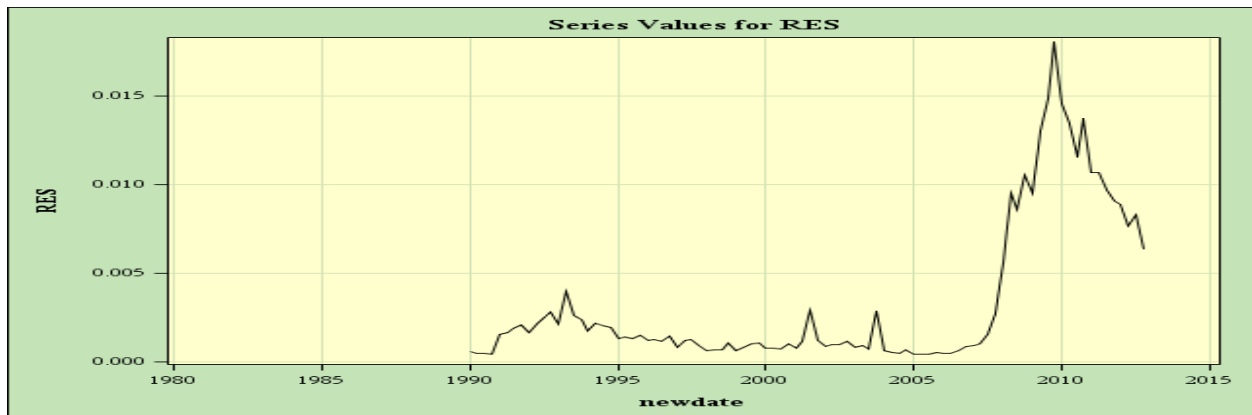
As BHCs assets recover, more and more banks are subject to Comprehensive Capital Analysis and Review (CCAR). In October 2012, the Federal Reserve finally ruled Covered Companies to disclose publicly the results of their stress tests under the Severely Adverse Stress Scenario, which describes the hypothetical evolution of certain specific macroeconomic and market variables consistent with a severely adverse post-war recession. And each covered company’s capital policy (stock buyback, dividend, etc.) is subject to the blessings on its CCAR from the Federal Reserve.

We can separate credit exposure of a loan into three components: PD (Probability of Default), LGD (Loss given Default) and EAD (Exposure at Default) and the expected loss is the product of these three risk parameters. Since EAD is the unpaid principal balance, we generally focus on PD and LGD. One method is to forecast “Net Charge Off” (NCO) which is the product of PD and LGD based on aggregated data.

METHOD ONE: USE SAS/ETS TO FORECAST NCO

We utilize SAS/ETS system to develop the models and making forecasts. The analysis performed by PROC ARIMA is divided into three stages, corresponding to the stages described by Box and Jenkins (1976): identification, estimation and forecast.

As an example, we download the excel file containing industry wide loan performance data maintained by FDIC from its website (<http://www2.fdic.gov/QBP/index.asp>). The FDIC quarterly banking profile stores performance history for banking industry loan portfolio. We use line 135 as residential mortgage (RES) up to 2002Q1 then switch to line 120.



Residential Mortgage NCO is a quarterly time series starting from the first quarter of 1990 and ending at the last quarter of 2012. Except for several small increases around 1993, 2001 and 2003, the net charge off is benign hovering around .1% until it jumped in 2009 and 2010 as a result of the great recession. And at the end of the data series, it stays elevated.

Here is the source code for the forecast:

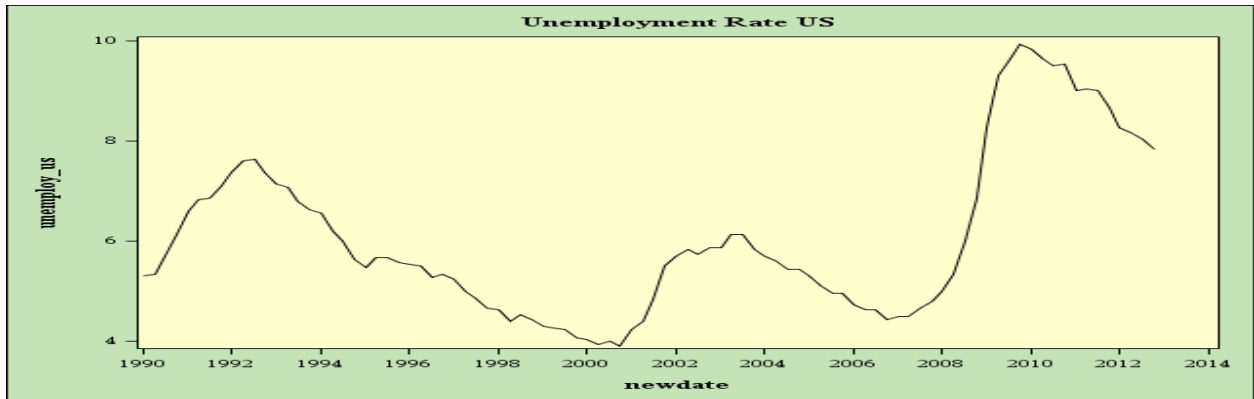
```
ods graphics on;  
proc arima data=nco_base plots=all;
```

```

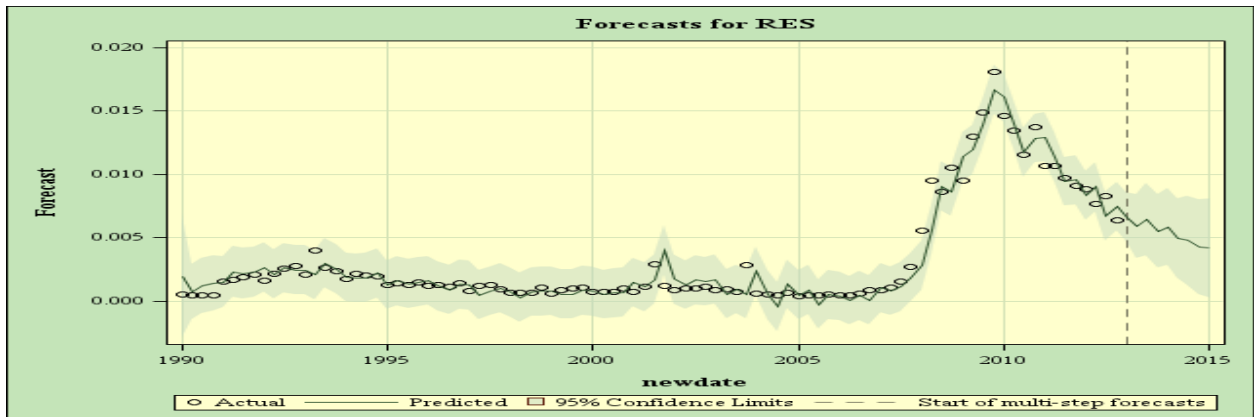
title "Forecasting RES baseline";
identify var=res cross=(unemploy_us)
minic scan esacf outcov=cov_res_base;
estimate p=(1,3,4) q=(6) input=(unemploy_us)method=ml;
forecast lead=9 id=newdate interval=quarter out=res_base;
run;
ods graphics off;

```

We combined built in model selection tools in SAS regression procedure with other model diagnostic techniques in SAS ARIMA procedure such as AIC, SBC, p value and so on to select the U.S. unemployment rate out of more than 25 plausible options. From the graph below, unemployment rate, although more volatile, move in sync with RES, with smaller increases in 1993, 2001, 2003 and a big jump around 2010.

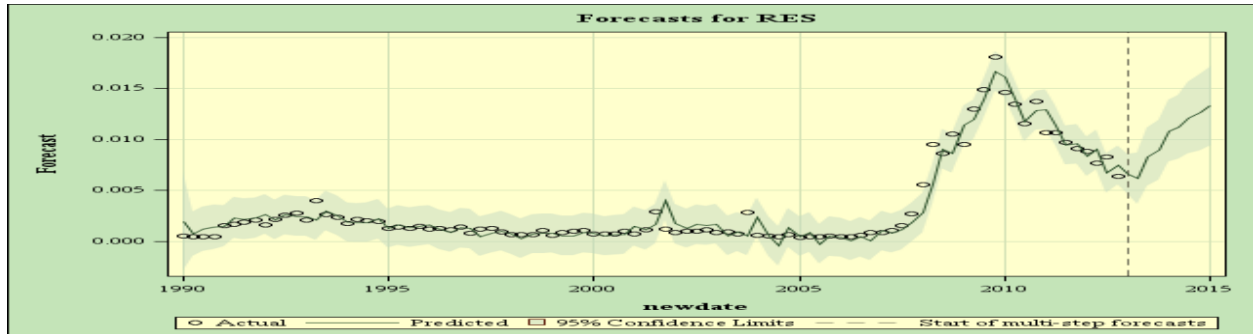


FORECASTING RESULTS UNDER BASELINE SCENARIO



Unemployment will decrease from its current elevated level in Moody's baseline scenario. As we would expect, NCO for RES under baseline scenario, will decrease but not to the pre-crisis complacent level any time soon.

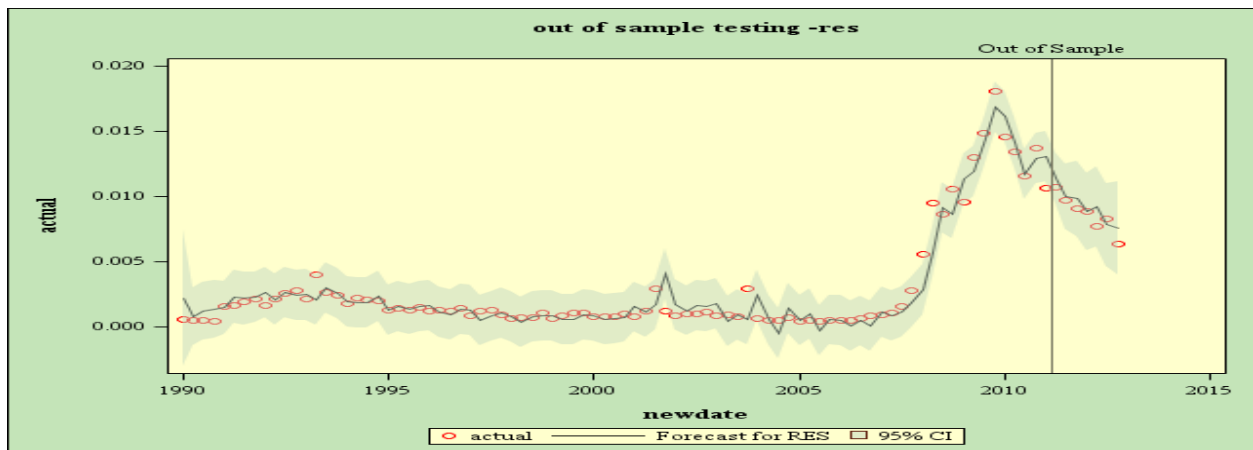
FORECASTING RESULTS UNDER STRESS SCENARIO



The stress scenario portrays a protracted slump economy in which unemployment rate will not only increase but also surpass the highest level in the aftermath of housing bubble. The model, intuitively sensibly, forecast a rising RES NCO under the stress scenario.

OUT OF SAMPLE TEST

Finally, we performed both the in and out of sample tests on the forecasting power of the model. For out of sample test, we preserved 7 periods at the end of the series beginning the 2nd quarter of 2011 and used the data before 2nd quarter 2011 to forecast the next 7 periods. The forecast from the 2nd quarter of 2011 to the last quarter of 2012 is then compared with the actual values. For both in and out of sample, we can see our forecast closely aligned with actual indicating a robust forecasting power of the model.



THE LIMITATION OF ARIMA

Auto-regressive integrated moving average (ARIMA) models are one of the most important and widely used linear time series models. We choose ARIMA model due to its statistical properties as well as the well-known Box–Jenkins methodology in the model building process. In addition, ARIMA has the advantages of accurate forecasting over a short period of time and ease of implementation. It has a track record of its forecasting accuracy and tends to exceed that of most time series models.

ARIMA models are quite flexible in that they can represent several different types of time series and implement various exponential smoothing models. ARIMA's major limitation is the pre-assumed linear form of the model. ARIMA models assume that future values of a time series have a linear relationship with current and past values as well as with white noise, so approximations by ARIMA models may not be adequate for complex nonlinear real-world problems. However, many researchers have argued that real world systems are often nonlinear (Zhang et al., 1998).

The Box-Jenkins approach involves three basic stages described in the above section. If the model developed in stages 2 and 3 does not meet expectations, the process is repeated and a new model is chosen and tested. To summarize, the main features of the Box-Jenkins model are:

- 1) It is quite complex in the model building process.
- 2) It out performs most other models in forecast accuracy in relatively short range.

- 3) It requires a relatively large amount of data (some authors feel at least sixty periods of data).
- 4) It is usually necessary to develop a new model whenever new data appear.

THE STRENGTH AND WEAKNESS OF LOAN LEVEL APPROACH

Loan-level model takes full advantage of the most granular level data available and the competing risk framework is conceptually sound. However, it usually involves with monstrous amount of data and requires massive computing power. Therefore, it is very difficult to develop, validate and implement the loan level model. In addition, some will criticize loan level model for its lack of transparency of transition of interim delinquency status.

BE AWARE OF ALL MODELS

Although it is trendy to use statistical models in financial risk management, we need to be aware of the limitations of all models. Models can't replace experiences, common sense and sound business judgment, can't catch the next black swan – the rare and extreme event, can't detect things not happened historically or simple not in the data, and most importantly, can't go beyond their underlying assumption, which, unfortunately, are many.

CONCLUSION

This paper provides an example using SAS ARIMA procedure to forecast loan portfolio loss as an essential part of the Federal Reserve's annual CCAR process. It also discusses the limitations and advantages in adopting the aggregate data approach. Besides, it also mentions the strength and weakness of a competing approach based on loan level data. At last, it warns against overly relying on models in financial risk management.

REFERENCES

- Box, G. E. P. and Jenkins, G. M. (1976), Time Series Analysis: Forecasting and Control, San Francisco: Holden-Day.
- Zhang G., Patuwo B.E., Hu M.Y. (1998), Forecasting with artificial neural networks: The state of the art; International Journal of Forecasting 14; 35– 62.

CONTACT INFORMATION

Your comments and questions are valued and encouraged. Contact the author at:

Name: Xinpeng "Tom" Wang
Enterprise: Huntington National Bank
Address: 5116 Shady Oak Drive
City, State ZIP: Hilliard, OH 43026
Work Phone: 614-480-5524
E-mail: xinpengwang@yahoo.com

SAS and all other SAS Institute Inc. product or service names are registered trademarks or trademarks of SAS Institute Inc. in the USA and other countries. ® indicates USA registration.

Other brand and product names are trademarks of their respective companies.