Incorporating External Economic Scenarios into Your CCAR Stress Testing Routines

Christian Macaro and Kenneth Sanford, SAS Institute Inc.

ABSTRACT

Since the financial crisis of 2008, banks and bank holding companies in the United States have faced increased regulation. One of the recent changes to these regulations is known as the Comprehensive Capital Analysis and Review (CCAR). At the core of these new regulations, specifically under the Dodd-Frank Wall Street Reform and Consumer Protection Act and the stress tests it mandates, are a series of "what-if" or "scenario analyses" requirements that involve a number of scenarios provided by the Federal Reserve. This paper proposes frequentist and Bayesian time series methods that solve this stress testing problem using a highly practical top-down approach. The paper focuses on the value of using univariate time series methods, as well as the methodology behind these models.

INTRODUCTION

The financial crisis of 2008 sparked a new wave of regulations for banks and bank holding companies (BHCs). One change was to augment the traditional economic capital modeling with macroeconomic scenarios as factors in simulations. These new regulations fall into two categories (Touryalai 2014): Comprehensive Capital Analysis and Review (CCAR) and the Dodd-Frank Act stress tests (DFAST). Adhering to these regulations requires a new method of forecasting the performance of banks and evaluating their ability to continue normal operations in a set of increasingly adverse economic scenarios (Hirtle 2012). This paper outlines two methods of incorporating these scenarios into projections of bank performance. The first method views the problem from a traditional, frequentist perspective. The second method uses a form of Bayesian model averaging to give the analysis tremendous power and intuition.

For banks, economic capital and operational risk modeling is a very complicated and resource-intensive activity. Traditional modeling of risk, or more appropriately risk simulation, involves modeling the probability of default (PD) and the loss given default (LGD) for each asset (Wang 2013). This paper introduces a novel way to approximate similar outcomes by using a top-down approach that has substantially lower data requirements and is far less computationally expensive. With this approach, these methods of incorporating the required economic scenarios can be extended to much smaller banks and BHCs that do not currently have such disaggregated or pristine data.

The models that are presented in this paper are specifically tailored to the problem of "too many" predictors in the context of estimating a model of losses. Three scenarios are provided by the Federal Reserve.

BASELINE, ADVERSE, AND SEVERELY ADVERSE ECONOMIC SCENARIOS

The most disruptive change to the existing process of oversight and compliance is the new requirement that external economic scenarios be built into economic stress tests. These scenarios, which are provided by the Federal Reserve (Fed), consist of 16 US macroeconomic indicators in three possible economic states: baseline, adverse, and severely adverse. Table 1 lists these macroeconomic indicators. The Fed recently expanded this list to include 12 additional international economic factors, though these are not part of the analysis in this paper.

Factor	Factor
Real GDP growth	10-year Treasury yield
Nominal GDP growth	BBB corporate yield

Factor	Factor
Real disposable income growth	Mortgage rate
Nominal disposable income growth	Prime rate
Unemployment rate	Dow Jones Total Stock Market Index (Level)
CPI inflation rate	House Price Index (Level)
3-month Treasury rate	Commercial Real Estate Price Index (Level)
5-year Treasury yield	Market Volatility Index (Level)

Table 1. The 16 Macroeconomic Indicators Provided by the Fed

Values for each of these variables are given for nine quarters into the future under three different economic situations. The Fed mentions specifically (Board of Governors of the Federal Reserve System 2014) that

> the baseline, adverse, and severely adverse scenarios include 28 variables, including economic activity, unemployment, exchange rates, prices, incomes, and interest rates. The adverse and severely adverse scenarios describe hypothetical sets of events designed to assess the strength of banking organizations and their resilience to adverse economic environments. They are not forecasts. The baseline scenario follows a similar profile to the average projections from surveys of economic forecasters. It does not represent the forecast of the Federal Reserve.

Figure 1 shows graphs of three of these variables—gross domestic product (GDP), GDP growth, and unemployment rate—to give you an appreciation for several of the scenarios. The new CCAR and DFAST regulations require these projections to be considered in forecasts of bank performance.
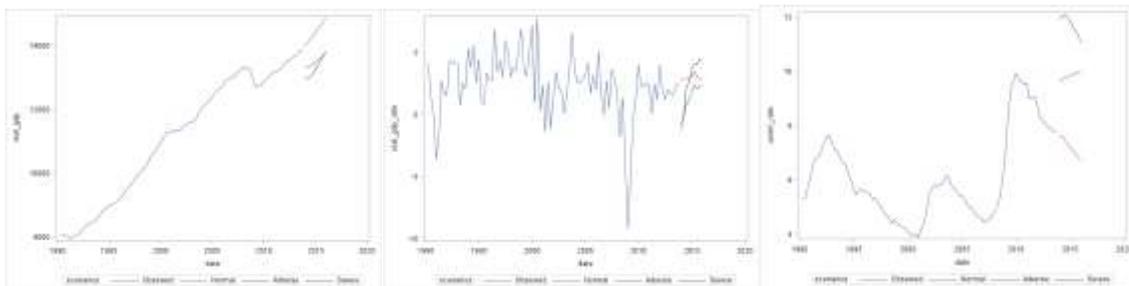

Figure 1. Graphs of GDP, GDP Growth, and Unemployment Rate Variables under the Three Scenarios

TOP-DOWN VS. BOTTOM-UP LOSS PROJECTIONS

The generally accepted method for simulating the losses and the reserve needs of banks and BHCs is to estimate models of PD and LGD for each asset in the portfolio. These models might have as arguments the current interest rate and borrower-specific information, among other factors. You arrive at a prediction of each asset's expected loss under a specific interest rate scenario by multiplying the two predicted values. Adding these values across various lines of business or asset classes then provides an estimate of the expected losses for certain accounts. Further perturbation of the arguments, such as the prevailing interest rate, allows for traditional stress tests to be performed. Experts generally consider this asset-level stress testing, the "bottom-up method," to be the preferred method for stress tests.

The bottom-up method is intuitively appealing because the performance of each asset is modeled directly, usually by a multinomial logit. You obtain estimates of the aggregate expected losses by adding the values of each asset. The results are very transparent. However, although the bottom-up method is usually the preferred method, its implementation presents considerable challenges. First, the bottom-up

method is data-intensive. In order to estimate and predict the likelihood of default for each asset, historical data are needed on each asset for the life of the asset. In large banks, where different assets belong to different divisions (think mortgages and commercial loans), or in BHCs, which might be even more fragmented, collecting loan information from each division is a monumental task. Even if these data are collected, the computational time to perturb and aggregate each asset is substantial, often many hours. For this reason, scenarios are often run overnight, which is appropriate for certain applications but not ideal for compliance officers who ask for assumptions to be challenged on the fly. The final challenge of the bottom-up method is the difficulty of extending these models with the economic scenarios. Again, each PD would need to have one or more economic scenarios as regressors. With macroeconomic data, which has zero cross-sectional variation and very little temporal variation, the likelihood that any effect could be accurately estimated is low. That is, how likely is it for the US unemployment rate or GDP to increase the probability that an individual will default on a mortgage? Finding any statistically significant effect is unlikely. And even so, which scenario should be included? Many banks and BHCs echo these criticisms of the bottom-up method. So what is the alternative? Consider a "top-down" approach.

The top-down approach starts with aggregate data, that is, the time series of losses for a given line of business. Rather than having to perturb regressors at the asset level, you perturb the regressors, or more accurately, the economic scenarios in aggregate. In this way, the average losses of the bank are directly related to the average values of economic scenarios. The data in this type of analysis are minimal, requiring only aggregate values. The statistical estimation is fairly routine, using linear regression or autoregressive techniques. And finally, the introduction of economic scenarios for forecasting future losses is computationally simple and straightforward. The trade-off is forfeiting the type of internal consistency provided by the bottom-up approach. This paper introduces an additional benefit of the top-down approach: it removes the need for variable selection on the economic scenarios. In other words, this approach lets you use all 16 macroeconomic variables.

TRADITIONAL APPROACH TO TOP-DOWN STRESS TESTING

The banking industry has explored several top-down methods as ways to incorporate the required economic scenarios. The traditional approach involves using a linear regression framework with contemporaneous and lagged values of one or more regressors to select a subset of variables in order to predict a particular time series. In essence, these steps are:

1. Regress the target variable (such as losses for an asset class) on one variable and its lags, choosing the coefficients that yield coefficients that are statistically significant by some measure. Retain these variables.

2. Repeat this process for all economic scenarios, maintaining the same target variable.

3. Combine all retained variables into a set of predictors, and re-estimate a model with all retained predictors from all prior regressions.

4. Forecast future values of the dependent variable under the three economic scenarios and generate required financial ratios.

This routine has been used by several banks on the Fed's CCAR/DFAST list and appears to have some acceptance among regulators.

There are two potential problems with this approach. First, the selection mechanism for each variable is somewhat arbitrary. That is, what justifies this iterative approach? A more natural selection mechanism for the process would be to include all regressors and lagged values of these regressors in one wide regression with the same target variable. This would ensure that the signs of the coefficients do not change from the first-stage regression to the second-stage regression, an undesirable change in the eyes of regulators. The second problem is that human judgment is used to select a subset of variables as predictors. This is required because of the dimensionality of the problem. That is, when you have a relatively short time series, and there are 16 predictors and lagged values of these predictors, you have a situation where there are many regressors and few observations. Variable selection is the typical method of solving this problem, though regulators have shown some apprehension about this approach. The next section outlines an alternative approach.

Bayesian Variable Selection

Linear regression has been the primary tool for top-down loss modeling. This statistical model is relatively simple, and it is commonly accepted as the basic tool for analyzing the relationship among several variables of interest. In the simplest scenario, the relationship between a variable of interest y and some covariates $x_1, \ldots, x_k$ is represented by

$$y_t = \beta_1 x_{1t} + \ldots + \beta_k x_{kt} + \epsilon_t$$

where $V(\epsilon_t) = \sigma$.

For example, if y represents the debts of a financial institution at time t and $x_1, \ldots, x_k$ represents a macroeconomic variable that summarizes the behavior of the economy over time, you might be able to predict the amount of debt affecting the financial institution if you know the future values of the macroeconomic variables. In this respect, the role of the federal regulator becomes crucial: macroeconomic forecasts of the economy are usually synthesized in terms of scenarios where the future values of $x_1, \ldots, x_k$ can be considered to be known.

But there is a problem: the ever-changing dynamic of the economy together with scarcity of data is usually incompatible with a large model, where the number of covariates k is large. Apart from not being feasible, a large model tends to perform well in the in-sample predictions but poorly in the out-of-sample predictions. These properties usually force practitioners into an extended analysis of which variable or covariate to include. As a result, some arbitrary choices need to be made, and the resulting model might not adequately represent reality.

To be more specific, if the pool of candidate covariates contains k components, there are $k\,(k+1)/2$ candidate models. The ideal goal is to select one single model, but in practice this is not possible, and using only one model might be frowned on by regulators. One way to overcome this problem is to abandon the idea that only one model is capable of representing reality and instead consider a scenario in which different models can emphasize different aspects of the same reality. In doing so, you need to evaluate the different models and their forecasts. There are many ways of scoring competing models; the Bayesian framework presented here is capable of assigning a probability to each individual model (Liang et al. 2008). A scoring system based on probabilities is ideal in many ways: it is simple to understand, optimal for decision theory, and extremely well suited to aggregation of the results.

If you have the set of $J = k\,(k+1)/2$ competing models $(M_1, \ldots, M_J)$ coming from the combinations of the k possible regressors, then the probability of each model can be evaluated as

$$P(M_i|y) = \frac{m_i p_i}{\sum_{j=1}^{J} m_j \pi_j}$$

where m_i and p_i are the marginal likelihood and the prior probability of model i. The evaluation of the marginal likelihood is crucial. It represents the likelihood of observing the data within a specific model after integrating out the uncertainty associated with the parameters of the model. In mathematical terms:

$$m_i = \int L_i(y|\beta)\pi_i(\beta)d\beta$$

where $L_i(y|\beta)$ and $\pi_i(\beta)$ are the likelihood function and the prior distribution for the parameters of the model M_i. In particular, each model M_i characterized by a subset of the regressors $x_1, \ldots, x_k$ is assigned a *g*-prior distribution of the form

$$\pi_i(\beta|\sigma^2, X) = N[\hat{\beta}, g\sigma^2(X^T X)^{-1}]$$

$$\pi_i(\sigma^2|X) \propto \sigma^{-2}$$

This set of conditional priors is conjugate in the sense that the analytical form of the posterior distribution is known. Moreover, the predictive posterior distribution and the Bayes factor are also known to have an

analytical form. The only problem is that the value of g is unknown and you need to come up with a guess: a common heuristic solution is to set g equal to the number of observations.

The following example analyzes data from the first quarter of 1990 to the third quarter of 2013. The macroeconomic variables are as follows:

- unemployment rate lags: 0, 1, 2

- House Price Index (HPI) lags: 0, 1, 2

- real GDP: 0, 1, 2

- real GDP rate: 0, 1, 2

- nominal GDP lags: 0, 1, 2

- nominal GDP rate lags: 0, 1, 2

Assuming that the intercept is a fixed parameter, you still have 262,144 possible models to explore. Exploring all the models is possible, but it seems impossible to decide which model to use. As an alternative, instead of focusing on a single model, focus on the most probable ones. The idea of associating probabilities with models is very powerful and very easy to interpret. For example, it is easy to see in Figure 2 that the number of regressors in this example should be approximately 6 or 7.
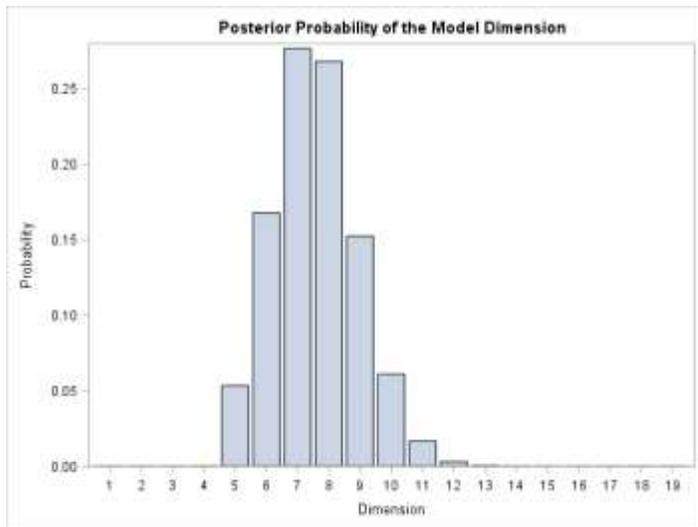

Figure 2. Probability Density Function of Number of Covariates in Models across All Specifications

You can find more information about which regressors should be included by looking at Table 2, which lists the variables included in the top nine most probable specifications.

Table 2 shows that unemployment rate and HPI are very important regressors, because contemporaneous values of these variables appear in all nine of the most probable specifications. Other variables that often appear in these models are HPI (2 lags), real GDP rate, and nominal GDP (2 lags).

Intercept	*	*	*	*	*	*	*	*	*
unem_rate	*	*	*	*	*	*	*	*	*
hpi	*	*	*	*	*	*	*	*	*
real_gdp									
real_gdp_rate				*		*		*	
nom_gdp	*							*	
nom_gdp_rate					*		*		*
unem_rate_1lag									
hpi_1lag									
real_gdp_1lag									
real_gdp_rate_1lag									
nom_gdp_1lag			*			*	*		
nom_gdp_rate_1lag									
unem_rate_2lag									
hpi_2lag	*	*	*	*	*	*	*	*	*
real_gdp_2lag									
real_gdp_rate_2lag									
nom_gdp_2lag		*		*	*				
nom_gdp_rate_2lag									
Probability	1.1%	1.1%	0.9%	0.8%	0.8%	0.6%	0.6%	0.6%	0.6%

Table 2. Regressors in the Nine Most Probable Specifications

A similar conclusion is obtained by looking at Table 3: the marginal probabilities of inclusion for unemployment rate and HPI are the highest (95% and 99%) among all the possible regressors.

Parameter	Probability of Inclusion	HPM	MPM
Intercept	100%	*	*
unem_rate	95%	*	*
hpi	99%	*	*
real_gdp	25%		
real_gdp_rate	36%		
nom_gdp	33%	*	
nom_gdp_rate	36%		
unem_rate_1lag	22%		
hpi_1lag	29%		
real_gdp_1lag	27%		
real_gdp_rate_1lag	15%		
nom_gdp_1lag	33%		
nom_gdp_rate_1lag	15%		
unem_rate_2lag	17%		
hpi_2lag	81%	*	*
real_gdp_2lag	22%		
real_gdp_rate_2lag	13%		
nom_gdp_2lag	36%		
nom_gdp_rate_2lag	13%		

HPM = highest probability model; MPM = median probability model.

Table 3. Probability That Regressor Should Be Included in Regression

Moreover, three hypothetical scenarios that span from the fourth quarter of 2013 to the fourth quarter of 2015 are assumed to be available: normal, adverse, and severe. By focusing on the 50 most probable models, you obtain 150 prediction paths. The graph in Figure 3 shows the predicted values for all three macroeconomic scenarios. The shading represents the probabilities associated with each forecast; red shading indicates the highest-probability models.
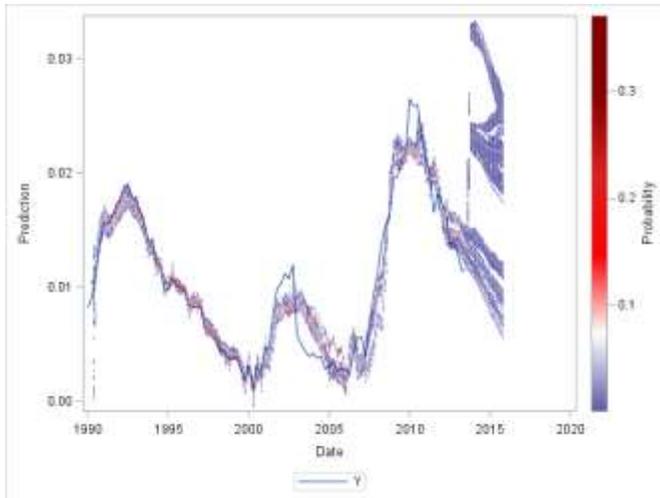
Figure 3. Prediction Paths of the 50 Most Probable Models in Three Different Scenarios

The three different scenarios (normal, adverse, and severe) clearly diverge in terms of the conclusions. The three graphs in Figure 4 show each scenario and set of forecasts separately.
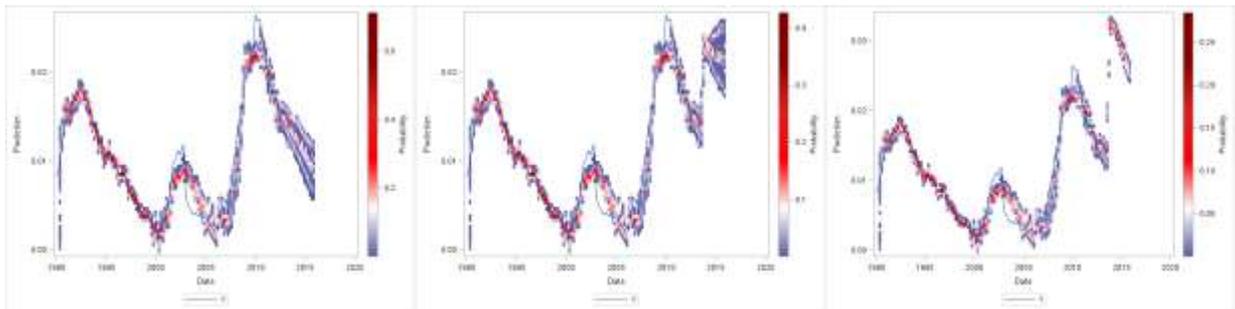

Figure 4. Prediction Paths of the 50 Most Probable Models in Detail

The temporal posterior distribution depicted in Figure 3 and Figure 4 can be difficult to view in the heat map representation. A common way of summarizing the huge amount of information is to focus on particular future date. Figure 5 shows the prediction distributions on April 1, 2014, January 1, 2015, and October 1, 2015. It is particularly interesting to notice that the three scenarios cluster in three distinct points for each date considered. These figures represent a cross section of predictions at a moment in time, representing a slice from the predictions seen in Figure 3 and Figure 4.
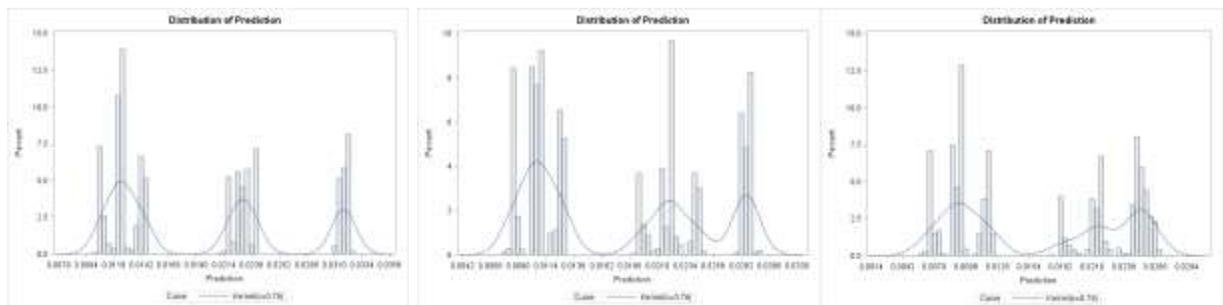

Figure 5. Prediction Distributions on April 1, 2014, January 1, 2015, and October 1, 2015 based on the 50 Most Probable Models across All Three Scenarios

8

CONCLUSION

Ever since the financial turmoil of the late 2000s, banks and bank holding companies have faced increased scrutiny by federal regulators. Of particular interest to the regulating bodies is the ability of banks to withstand prolonged, severe economic recessions. To address the need for banks to evaluate their economic capital positions under stressed economic conditions, several pieces of legislation were enacted. Among these were a series of requirements that banks must consider three different macroeconomic environments in their forecasts. The scenarios themselves are provided by the Federal Reserve as part of a set of 16 US economic indicators plus 12 international economic indicators. Banks face the challenge of creating models that use this information.

This paper proposes a new method for banks to use in incorporating the three economic scenarios into their stress test modeling. It discusses the prevailing top-down approach to incorporating these scenarios and provides a new, less arbitrary approach. This new approach should be more acceptable to regulators because it includes all macroeconomic variables in the analysis and gives them some weight in the forecasting of future values of losses. This method scales to more macroeconomic scenarios and provides a mechanism for quickly predicting losses under different economic scenarios.

REFERENCES

Board of Governors of the Federal Reserve System. (2014). "Stress Tests and Capital Planning." October 23. http://www.federalreserve.gov/newsevents/press/bcreg/20141023a.htm.

Hirtle, B. (2012). "CCAR 2012: Overview and Stress Scenario Results." Federal Reserve Bank of New York. Accessed January 12, 2015. http://www.newyorkfed.org/education/pdf/2012/Hirtle_comprehensive_capital_analysis_review.pdf.

Liang, F., Paulo, R., Molina, G., Clyde, M. A., and Berger, J. O. (2008). "Mixtures of *g* Priors for Bayesian Variable Selection." *Journal of the American Statistical Association* 103:410–423.

Touryalai, H. (2014). "Wall Street's March Madness: What to Know about Bank Stress Tests." *Forbes*. http://www.forbes.com/sites/halahtouryalai/2014/03/19/wall-streets-march-madness-what-to-know-about-bank-stress-tests/.

Wang, X. (2013). "Use SAS/ETS to Forecast Credit Loss for CCAR." *Proceedings of the Midwest SAS Users Group 2013 Conference*. Cary, NC: SAS Institute Inc. Accessed January 8, 2015. http://www.mwsug.org/proceedings/2013/FS/MWSUG-2013-FS03.pdf.

ACKNOWLEDGMENT

CONTACT INFORMATION

Your comments and questions are valued and encouraged. Contact the authors:

Kenneth Sanford
SAS Institute Inc.
859-227-9787
Kenneth.Sanford@sas.com

Christian Macaro
SAS Institute Inc.
Christian.Macaro@sas.com