Stress Testing with the Help of Bayes' Theorem

Notes from the Vault

Mark J. Jensen
February 2016

This year's Federal Reserve stress test scenario includes short-term Treasury rates declining to previously unrealized levels of negative 50 basis points. Because the Fed expects banks to use quantitative models to project their losses, revenues, expenses, and capital levels for this scenario, not having a history of negative interest rates makes modeling difficult and forecasts unreliable.

This post describes how Bayes' theorem can be used to overcome the difficulty associated with modeling and forecasting when the stress test scenario is unlike a bank's past experience. It begins by describing the expectations the Federal Reserve has for how stress test forecasts are to be calculated. Next, I present and explain Bayes' theorem. I then apply Bayes' theorem to a case where historical data are limited and show how the data from other banks operating under a negative rate environment can be used to update a bank's probability model. By learning from a cross-section of data from banks operating in an economic environment similar to the stress test scenario, Bayesian probability models can be more reliable and have less uncertainty associated with their point forecasts.

Federal Reserve stress tests

Each year the Federal Reserve is required by law to conduct a stress test of every bank holding company (BHC) with over $50 billion in assets (12 CFR 225.8). Called the Federal Reserve's annual Comprehensive Capital Analysis and Review (CCAR), it requires these BHCs to project losses, revenues, expenses, and capital levels nine quarters into the future under the conditions specified by three hypothetical economic and financial scenarios. If a BHC can hypothetically meet all regulatory capital requirements over the nine-quarter horizon, it quantitatively passes the Fed's stress test.

These hypothetical economic scenarios consist of a baseline, adverse, and severely adverse scenario. The scenarios range from a future economic environment similar to the consensus forecast for the U.S. economy as projected by professional forecasters for the baseline case to a dramatic downturn in the economy and financial markets for the adverse and severely adverse scenarios.

In the first couple years of the stress test, the Federal Reserve suite of adverse and severely adverse scenarios closely resembled the economic conditions of the Great Recession, if not in magnitude, at least in direction. However, this and last year's scenarios are more hypothetical and have economic and financial conditions not previously experienced in the United States. For example, last year's adverse scenario had future interest rates increasing while at the same time output was declining and unemployment was increasing. In this year's severely adverse scenario, the three-month Treasury rate drops to negative 50 basis points and stays there for the rest of the nine-quarter stress test time horizon.

Firms are expected to use quantitative models when possible for forecasting. A bank holding company's quantitative model needs to be sensitive to the macroeconomic factors of the scenarios and to other risk drivers and are preferred to qualitative approaches to forecasting unknowns. Granular, firm-specific data should be used to estimate and fit the models. If the necessary data are lacking, or there are known limitations to a model, the Federal Reserve is open to the firm using qualitative model overlays and expert judgment to overcome such model and data shortcomings (Federal Reserve SR 15-18, page 23).

Probability models

Probability models are often used to quantify one's uncertainty about the future. As abstractions of reality, probability models help to refine and organize our understanding of the real world and increase the reliability—and quantify the uncertainty—around a prediction. Based on how well a probability model has performed in the past, it can change and improve. Poor past forecasting performance leads to the replacement of existing models with a new model, while small predictions errors and increasing precision around the forecast result in a model being enhanced and refined to deal with those situations over which it performed poorly.

The success of a probability model for stress testing purposes depends on the availability of empirical data representative of the conditions under which the model forecast will be conditioned on. However, it is important to remember that "all models are wrong, but some are useful" (Box, 1979). Hence, expert judgment always plays a role in modeling and should be used in judging its success, especially when the expert expects a major change from the past.

A model's empirical success depends on how good the data used to estimate it are. For centuries, statisticians have struggled with not having the necessary data needed to make informed predictions. Accurate model-based stress test predictions would ideally include significant experience under conditions similar to the stress test scenario. Fortunately, severely adverse scenarios are rare in real life, but that implies stress test models are being asked to make predictions under less than ideal conditions for the model.

Data limitations extend beyond not having a history that resembles the stress test's hypothetical scenarios. It also includes flawed data or the lack of quality data. According to a report by the Senior Supervisors Group, "the area of greatest concern remains firms' inability to consistently produce high-quality data" (Senior Supervisors Group, page 1). Until the financial crisis, BHCs committed limited resources to data collection and to the management of their information systems. As a result, good-quality, loan level information is often only found for the postcrisis period.

Uncertainty in forecasts

One would hope a BHC's level of uncertainty around its model's conditional forecast would decline as the hypothetical scenario looked more and more like its past. Mandatory elements of a BHC's capital plan, however, only require point forecasts of revenues, losses, reserves, and capital levels over the nine-quarter planning horizon under the three different economic scenarios (Federal Reserve Board, page 5). Neither the firms, nor the Federal Reserve, are required to report the level of uncertainty around their point forecasts.

The Federal Reserve instead asks firms to understand the sensitivity of these forecasts to changes in model inputs and assumptions (, page 20). For example, sensitivity analysis can consist of a firm taking its estimated model and projecting revenues, losses, expenses, and capital levels under a variety of different economic scenarios to see how the model forecasts perform under normal and stressful conditions. Although the possible scenarios are numerous and of a wide variety of stress levels, the model's underlying parameters are held fixed to those values that "best" fit the firm's historical experience. This approach ignores the parameter uncertainty associated with the different scenarios. Parameter values would likely be different if the historical data matched or looked similar to the hypothetical economic scenarios, and the projected results would likely look different, too.

Bayesian analysis

Fortunately, advances have been made that help quantify the uncertainty around a model's prediction, especially when the prediction is made under conditions different from past historical experiences or when the data used are questionable or unreliable. These advances in measuring the uncertainty around future events rely on applying a 250-year-old probability theorem attributed to the Reverend Thomas Bayes.

Bayes' theorem amounts to the action of moving from an initial probability statement about one's current state of knowledge about how the world works to a new probability statement as new information is observed. For example, suppose XYZ Corporation's earnings are to be forecasted for the next quarter and I currently expect its earnings to increase by at least 5 percent. Suppose before XYZ announces its earnings, ABC Corporation announces its earnings increased by 10 percent. Bayes' theorem formally shows how I would update my initial expectations for XYZ's earnings by how informative ABC's earnings are to those earnings.

Our initial understanding about the probability of the future event E occurring—for example, next quarter's earnings for XYZ Corporation being at least 5 percent—can be abstractly represented by $p(E)$. The likelihood the event E led to the empirical observation D, for example, earnings of at least 5 percent for XYZ having led to ABC announcing earnings of 10 percent, can be denoted by $p(E|D)$. Then by Bayes' theorem

$$p(E|D) = \frac{p(E)p(D|E)}{p(D)}$$

In the above expression, empirical information contained in $p(D|E)$ is combined with the initial probability model for E, $p(E)$. By combining the new empirical information of D with the current probability model for E, one's probability model about E gets updated to $p(E|D)$. As a probability model, $p(E|D)$ is an expression that answers what are the odds of E occurring, along with what is the probability of E taking on a variety of different values. Hence, $p(E|D)$ is a probability statement about one's uncertainty about the outcome of E, given the data D.

If the information contained in the data D is completely irrelevant to the future event E, Bayes' theorem delivers $p(E|D) = p(E)$. In other words, no learning occurs when the data D is irrelevant to the event E, and the best one can do is continue to rely on his or her initial probability model $p(E)$. This would be the case in the example if XYZ and ABC are completely unrelated companies, for instance, if they are firms from entirely different sectors of the economy. Under these circumstances, it also holds that $p(D|E) = p(D)$, meaning the occurrence of the data D does not depend on E.

Cross-sectional information

Bayes' theorem is useful when a prediction is made conditional on extreme events like those required under the Federal Reserve's annual stress test. For example, suppose E_1 represents the future losses of a firm for a particular stress test scenario, S. The firm's initial probability statement about its future losses is based on past economic conditions, S_0. This initial probability statement is represented by the probability model, $p(E_1|S_0)$. When the BHC incorporates its bank level data, D_1, into its probability statement, it applies Bayes' theorem and updates its model to

$$p(E_1|D_1,S_0) = \frac{p(E_1|S_0)\,p(D_1|E_1,S_0)}{p(D_1|S_0)},$$

where the updated model is conditional on D_1 and past economic conditions, S_0. If the economic conditions of the stress test scenario, S, are dramatically different from S_0, forecasts with this updated model will be unreliable, and the level of reliability will decrease the more S differs from S_0. Qualitative adjustments or model overlays will need to be applied to project losses, which are projections based on the probability statement $p(E_1|S)$.

Now suppose the firm is able to pool information from across the globe or from other banks about losses that occurred under conditions like those in the stress test scenario, S. In our original example about predicting XYZ Corporation's earnings, this pooled information would be ABC Corporation's recent announcement of 10 percent earnings under the conditions XYZ is currently experiencing but for which it did not have prior experience.

For this year's severely adverse scenario case of negative short-term interest rates, this information pool could include banks from Switzerland, Japan, Denmark, and Sweden—all economies that have experienced negative interest rates. Letting D_2, D_3, D_4, and D_5 represent the historical cross-section of bank data for firms from Switzerland, Japan, Denmark, and Sweden, respectively, our original bank updates its probability model for losses when S is assumed to occur via Bayes' theorem to

$$p(E_1|D_2,D_3,D_4,D_5,S) = \frac{p(E_1|S)\,p(D_2,D_3,D_4,D_5|E_1,S)}{p(D_2,D_3,D_4,D_5|S)}.$$

If the pooled data are informative about the future losses E_1 under the scenario S, in other words, the likelihood that $p(D_2, D_3,D_4,D_5|E_1,S)$ is large, then the firm can be more confident about the potential for losses under S than it was when it did not use the cross-section of bank data. Using bank data from countries that experienced negative interest rates, the firm no longer needs to rely on its experts or place overlays on its models when forecasting E_1.

For a firm whose data did have bearing on the losses under the scenario S, it could update its current probability statement $p(E_1|D_2,D_3,D_4,D_4,S)$ with the information found in $p(D_1|E_1,S)$ and by Bayes' theorem arrive at the even more informed probability statement

$$p(E_1|D_1,D_2,D_3,D_4,D_5,S) = \frac{p(E_1|D_2,D_3,D_4,D_5,S)\,p(D_1|E_1,S)}{p(D_1|S)}.$$

This example shows how using Bayes' theorem can identify how losses are distributed across the population of banks by observing a cross-section of banks that have operated under conditions similar to the stress case scenario. This pooling also applies to learning from the losses incurred by other banks that have failed, such as using data from Lehman Brothers, Wachovia, and Washington Mutual to make more informed predictions about losses under stress. When data from the failed banks are excluded, a firm's prediction for E_1 will naturally suffer from survivorship bias, since $p(E_1|D_1,D_2,D_3,D_4,D_5,S)$ uses data from banks that survived the financial crisis.

So why isn't Bayes' theorem used more widely, such as in the annual stress test? Although Bayes' theorem has been around since the 18th century, it has only been since the 1990s that computers and algorithms for sampling from multivariate probability distributions have overcome the difficult analytical calculations originally required by Bayesian analysis. However, around the time Bayes' theorem became practical, banks had already adopted the non-Bayesian risk management tools that underlie current stress testing procedures.

Conclusion

Bayesian analysis has considerable potential as a tool for improving bank stress tests. Its use would give the Federal Reserve and bank holding companies the ability accurately to project stress test losses and properly account for the uncertainty in their forecasts. Bayes' theorem also leverages the information found in the cross-section of the entire population of firms. The ability to use cross-sectional data reduces the uncertainty around a firm's forecast of revenues, losses, expenses, and capital when the economic and financial conditions underlying the forecast are extreme. Given the benefits of this new statistical tool but very old statistical idea, model developers and users involved with stress testing should familiarize themselves with Bayesian methods.

Mark J. Jensen is a financial economist and policy adviser at the Atlanta Fed. The author thanks Mark Fisher, Kris Gerardi, and Larry Wall for helpful comments on the paper. The view expressed here are the author's and not necessarily those of the Federal Reserve Bank of Atlanta or the Federal Reserve System. If you wish to comment on this post, please email **atl.nftv.mailbox@atl.frb.org.**

References

Box, G.E.P. (1979). "Robustness in the Strategy of Scientific Model Building," in Launer, R.L., Wilkinson, G.N., *Robustness in Statistics* (New York: Academic Press), 201–236.

RELATED LINKS ON OTHER SITES: Mark Jensen's web page   • Basel III and Stress Tests   • Bayes' Theorem for Everyone

Careers