MOODY'S
ANALYTICS

WHITEPAPER

Peer-Group Analysis in Bank Call Report Forecasts

Prepared by

Brian Poi
Brian.Poi@moodys.com
Director

Contact Us
Email
help@economy.com
U.S./Canada
+1.866.275.3266
EMEA
+44.20.7772.5454 (London)
+420.224.222.929 (Prague)
Asia/Pacific
+852.3551.3077
All Others
+1.610.235.5299
Web
www.economy.com
www.moodysanalytics.com

Introduction

Banks often build forecast models relying solely on internal data and a few dozen macroeconomic variables, but disentangling the effects of industry trends and bank-specific decisions is nearly impossible with limited data. As a result, these models cannot forecast accurately.

The Moody's Analytics Bank Call Report Forecasts database provides forecasts for more than 200 balance sheet and income statement variables at the industry level. We leverage those industry forecasts to produce accurate peer-group and bank-specific forecasts. Our methodology is consistently applied to all 6,100 banks in the U.S., allowing analysts to make fair comparisons among banks.

Peer-Group Analysis in Bank Call Report Forecasts

BY BRIAN POI

Banks often build forecast models relying solely on internal data and a few dozen macroeconomic variables, but disentangling the effects of industry trends and bank-specific decisions is nearly impossible with limited data. As a result, these models cannot forecast accurately.

The Moody's Analytics Bank Call Report Forecasts (CRF) database provides forecasts for more than 200 balance sheet and income statement variables at the industry level. We leverage those industry forecasts to produce accurate peer-group and bank-specific forecasts. Our methodology is consistently applied to all 6,100 banks in the U.S., allowing analysts to make fair comparisons among banks.

1. Introduction

The Moody's Analytics Bank Call Report Forecasts database provides scenario-specific forecasts for more than 200 balance sheet and income statement variables at the industry level. Mehra (2016) describes the methodology underlying the CRF.[1] Hughes and Poi (2016) illustrate the power of peer-group analysis in stress-testing and strategic planning activities.[2] This paper provides the details on how we use those industry-wide forecasts to produce more accurate forecasts for peer groups and individual banks.

The remainder of this section provides an overview of the challenges banks face in forecasting the line items on their financial statements. Section 2 describes the effect of mergers and acquisitions on internal bank data and how we remedy the situation. Section 3 provides an overview of how we relate bank-specific data to the industry data in the Bank CRF. Section 4 discusses how we can create custom peer groups as an alternative to comparing a bank with the entire industry. Sections 5 and 6 are more technical. Section 5 describes our automated forecast tool that selects among dozens of candidate models for each variable, and section 6 provides benchmarking results establishing its validity. Section 7 briefly concludes.

A bank relying solely on internal performance data and the standard macroeconomic variables released each year by the Federal Reserve as part of the Comprehensive Capital Analysis and Review exercise faces formidable challenges in forecasting individual line items on its financial statements. The historical data are often limited, covering 10 years or less. In addition, the bank may have bought other banks in the past, so its internal data show discontinuities that must be handled in the forecast models. Finally, the models must control for not just macroeconomic factors but also internal management decisions that may have an effect on the variable the model forecasts.

We propose a simple, coherent methodology that allows us to forecast and stress-test a bank's entire balance sheet and income statement. Our methodology is applied consistently to all 6,100 banks in the U.S., and our forecasts are more robust and more accurate than those built solely on internal bank data.

Passed in response to the financial crisis, the Dodd-Frank Act requires large and mid-size banks to develop models of asset and liability balances, loan originations, interest and noninterest income and expenses, and credit losses. These models are combined to forecast the capital position of the institution under various economic scenarios. Although the procedures are not required of smaller institutions, those banks could certainly benefit from them if the tests could be done quickly and with limited resources.

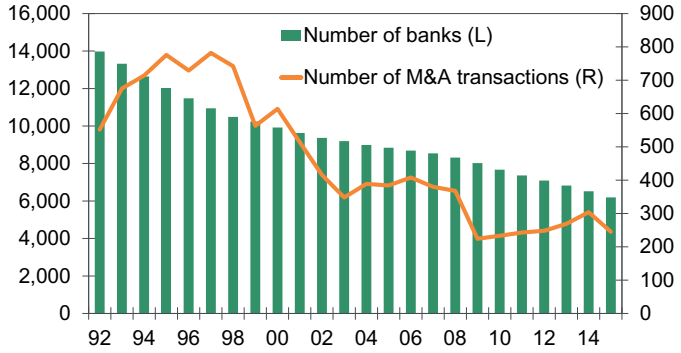
Banks' stress tests have primarily used complex bottom-up techniques to model a specific, narrowly defined cash flow or credit metric. The modelers source data for that line item from the bank's internal data warehouse and then build a model that relates that variable to macroeconomic factors. Once hundreds of models have been developed, risk managers combine the forecasts from those models to calculate the capital position of the bank under various economic scenarios.

Despite the best of intentions, such an approach to stress-testing is far from ideal. The forecasts from the hundreds of individual models may not be consistent with one another. Even as banks have invested in computational infrastructure, running these types of stress tests can take weeks. In an ideal world, bank executives would conceive

1 S. Mehra, "Moody's Analytics Bank Call Report Forecasts Methodology" (Forthcoming, 2016).

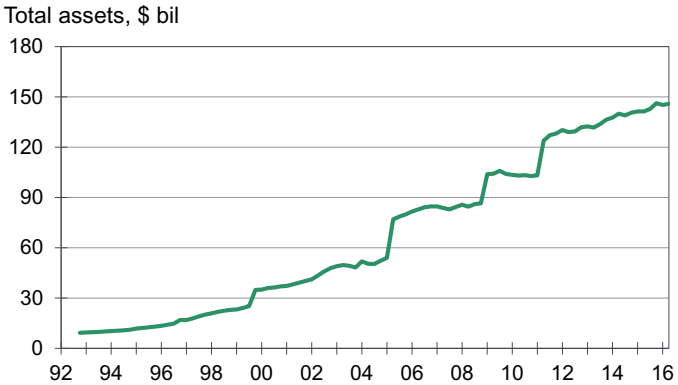
2 A. Hughes and B. Poi, "Stress-Testing and Strategic Analysis Using Peer Analysis," *Moody's Analytics Risk Perspectives, 8th ed.* (2016): 61—70.

Chart 1: Banking Consolidation Continues



Sources: FDIC Statistics on Depository Institutions, Chicago Fed, Moody's Analytics

Chart 2: M&A Activity Distorts Internal Data



Sources: FDIC Statistics on Depository Institutions, Moody's Analytics

of an economic scenario in the morning, have results from the risk management group by midafternoon, and have a contingency plan in place before quitting time.

Our methodology can be applied consistently to all FDIC-insured banks. Moreover, because we use publicly available data, we can forecast not just a particular bank but that bank's competitors and acquisition targets as well.

2. Mergers and acquisitions

In the last 25 years the number of banks has declined by more than half, from nearly 14,000 at the end of 1992 to about 6,100 today. Most of that reduction is attributable to mergers and acquisitions. In Chart 1 we plot the number of banks as well as the number of mergers and acquisitions transactions recorded in the Federal Reserve Bank of Chicago's database of M&A activity.[3]

M&A activity does not play a role in the industry-level aggregate data because a transaction simply represents a shift in assets and liabilities from one FDIC-insured institution to another. For the modelers at a bank trying to build forecasts based solely on internal data, M&A activity erects formidable roadblocks.

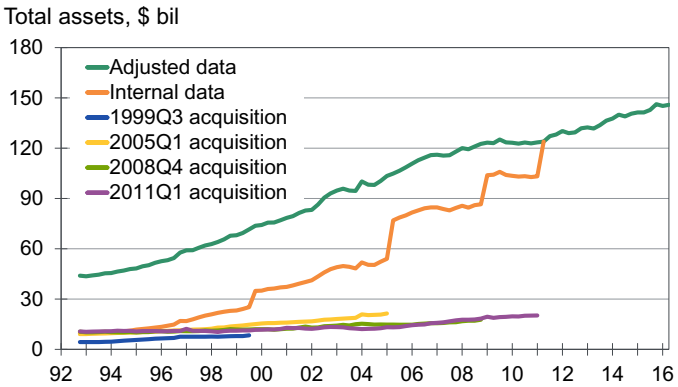
Chart 2 illustrates the problem. Our hypothetical bank acquired four banks, in 1999Q3, 2005Q1, 2008Q4 and 2011Q1. The modeler obtains and plots the internal data, resulting in the unadjusted series in our chart. One might argue that the bank will also have all the historical data for the banks it acquired, but in practice we have found that not to be typical. Many times the acquired bank's IT infrastructure is different from the acquiring bank's. Then the acquirer focuses on updating current data but does not update its historical records. In other cases, we have worked with modelers at banks who claim to have archived data from acquired banks only to find out when the modelers ask for those data that they are not retrievable. Particularly before the era of regulatory stress-testing, bank managers often decided that the costs of fully integrating acquired banks' historical data into their own systems were unjustified.

Chart 2 extends back to 1992Q4, the start of the FDIC's Statistics on Depository Institutions (SDI) database from which we obtain the call reports. For most banks the data available to modelers are much shorter. Say our bank's modelers are more fortunate than most, with data extending back to 2001. Using data through 2014 to build models (preserving 2015 and 2016 data for forecast evaluation) gives the modelers 56 observations for estimation. The models must control for the discontinuities due to the three acquisitions that have occurred over the estimation sample. Most risk models use indicator variables to control for these events.

Using Harrell's[4] (2001) more permissive rule of thumb of needing at least 10 observations per regressor in a linear regression model implies a model for our bank can have no more than two or three macroeconomic factors once we include the three indicator variables. Using Harrell's more conservative rule of 20 observations per regressor implies that if we control for the acquisitions we cannot control for the macroeconomy.

Our solution is to combine bank-level call report data from the FDIC SDI with the merger database maintained by the Chicago Fed to create series that reflect what the bank we are modeling would look like had it always owned the banks it has acquired (see Chart 3). The series marked "Internal data" is

Chart 3: Adjusting Data for Mergers



Sources: FDIC Statistics on Depository Institutions, Moody's Analytics

3 We have excluded about 1,100 acquisitions by government agencies that represent bank failures.

4 F. E. Harrell, *Regression Modeling Strategies With Applications to Linear Models, Logistic Regression, and Survival Analysis* (New York: Springer-Verlag, 2001), 61.

the same as in Chart 2; it represents the data we obtained by simply extracting that bank's data from the quarterly call reports. Our bank acquired a small bank in 1999Q3. Having the historical call report data allows us to obtain that small bank's data prior to the acquisition date, also shown in Chart 3. Similarly, we can obtain the pre-acquisition data for the other three banks our bank acquired.

Our merger adjustment procedure starts with the bank's internal data and proceeds chronologically across mergers to construct the final series. We combine the internal data with the data for the 1999Q3 acquisition. We then combine that with the data for the 2005Q1 acquisition and so on. When we are done, we obtain the adjusted data that represent our bank as if it had always owned the other banks.

Our example bank acquired the other banks outright, as is most often the case. Occasionally, though, an institution may acquire only part of another bank, perhaps the branches located in a particular area. The Chicago Fed merger database records these partial mergers and acquisitions, and the merger adjustment program we use handles these events as well.

The merger database records partial acquisitions, but it does not record the fraction of the selling bank that has been sold. Nevertheless, we can infer it from the data. Suppose bank A agrees to acquire part of bank B on a particular date. Then bank B's quarterly data exist in the call report database both before and after the acquisition date. We look at bank B's total assets in the quarter just after the sale and divide it by the bank's total assets in the quarter prior to the sale to obtain the scale of the acquisition. We then adjust all the line items in the call report by this scale.

Transactions representing asset transfers between banks and nonbank entities are not included in the Chicago Fed mergers database, but those transactions are rare. For the vast majority of banks we can produce call report data that is free from discontinuities attributable to M&A activity.

Adjusting data for M&A removes jumps in the data, but it may not eliminate all regime shifts. Say bank A acquired bank B, but prior to the acquisition, banks A and B were run by managers with different risk preferences or strategies for growth. Variation in our adjusted series after the acquisition reflects bank A's management style, but the adjusted data before the acquisition reflect a combination of bank A's and bank B's management styles. If bank B's managers reacted differently to macroeconomic factors than bank A's managers, then regression models of the adjusted data would not provide true measures of how the series will respond to macroeconomic conditions in the future. There are mitigating factors that suggest this source of potential bias is limited, however. In most acquisitions, the acquiring bank is much larger than the bank being acquired, so in the combined, adjusted data the effect of bank B's management will be relatively small. Moreover, although different banks may have different strategies, all banks and their customers respond to the same underlying economic conditions. Thus, the degree to which bank B's sensitivity to macroeconomic conditions differs from that of bank A's is arguably minor. Overall, the benefits of adjusting for M&A activity more than justify the behavioral assumptions we must make to use it.

3. Using industry data to forecast a bank

Given a set of scenario-specific forecasts for the banking industry from the CRF, our task is to produce forecasts for an individual bank that reflect industry trends but at the same time account for the bank's own strategy. In most industries, a firm forecasts its sales volume and other performance metrics by first considering the overall industry and then determining how much market share it can expect to capture. If an inventor develops a brand new widget and wants to predict its potential sales, she first looks at the market for all makes and models of widgets. Once she knows the size of the entire market, then she can speculate on the share of the market that will embrace her new widget and forecast her sales. We approach the items on a bank's balance sheet the same way.

Suppose we want to forecast a bank's stock of commercial real estate loans held on its book. Then for each quarter we define the bank's market share of CRE loans as its CRE balance divided by the industry-level aggregate balance of CRE loans reported in the CRF. In general, forecasting a market share turns out to be easier than forecasting a dollar amount. With an industry-level forecast and a bank-level market share forecast in hand, forecasting the bank's dollar balance of CRE loans is trivial.

For many line items on a bank's financial statement, we can interpret a bank's market share over the business cycle as a measure of the bank's preference for safer versus riskier (yet potentially more profitable) loans. If a bank has been neither aggressive nor conservative in its CRE lending, then its market share will be relatively constant. A conservative bank may have a countercyclical market share. As the market for CRE projects becomes hot, the conservative bank's market share may contract as it chooses not to pursue riskier, higher-leverage loans. When the CRE market enters a downturn, more aggressive banks will cut back lending and focus their efforts on preserving capital and nursing nonperforming loans back to health, reducing their market share. The conservative bank, not having lent to the risky borrowers, is in a position to lend to safer projects that come along, increasing its market share.

We identify five trajectories for market shares. Most market shares are relatively flat, as most banks tend to follow broader industry trends. Some market shares rise over time as those banks pursue strategies to grow faster than the industry. Other market shares fall over time as, for whatever reason, those banks' managers do not pursue growth strategies. Aggressive banks will have procyclical market shares, while conservative banks will have countercyclical market shares. An individual bank may shift strategies over time, causing structural changes in the evolution of its market shares.

We also distinguish among three determinants of a bank's market share. Macroeconomic factors can affect market share; as we just explained, a bank's risk appetite can induce cyclical behavior in its observed share. Market shares may also have a trend component; for example, a bank located in

a faster-growing region of the country will exhibit an upward trend in its market share vis-à-vis the industry. Finally, a bank's market share is affected by idiosyncratic factors peculiar to a bank. These factors include things such as marketing campaigns, pricing decisions, and other strategic actions the bank's managers take independent of broad macroeconomic conditions.

Developers who build models solely on internal data must control for trends, macroeconomic factors and idiosyncratic events. Given the limited span of data many banks have, that is an inordinate task. Moreover, a model that controls for macroeconomic factors but not idiosyncratic events is misspecified, so forecasts under alternative economic scenarios will be inaccurate.

By focusing on market shares, our methodology subsumes the need to control for macroeconomic and industry-wide factors. We can instead focus on modeling the bank's risk appetite and idiosyncratic factors to the extent they correlate with macroeconomic variables. Bank managers react to economic conditions, so the idiosyncratic component of market share will exhibit some correlation with the macroeconomy. We assume the portion of the idiosyncratic component that is not correlated with the macroeconomy is constant or evolves only slowly over time. That assumption implies the constant term in our regression models will adequately control for that portion of the idiosyncratic component.

The market share approach has intuitive appeal for balance sheet items such as loan balances and deposits, but it does have one shortcoming. The share approach does not work if the variable being forecast can be zero or negative. Suppose we are forecasting noninterest income. If the industry as a whole reports positive noninterest income for a quarter but a particular bank reports a loss, its share of industry noninterest income will be negative. We were initially tempted to simply let shares be negative, build forecast models that do not constrain shares to be positive, and use our share approach for these variables. However, our testing showed that such forecasts once converted to dollar amounts invariably performed poorly with counterintuitive trajectories under alternative scenarios.

For those variables we use what we term a "beta" model analogous to the beta parameter in the well-known capital asset pricing model. Here we model the bank-level variable as a function of the industry-level analogue as well as other macroeconomic variables. Changes in the industry aggregate drive the bank-level variable, and the macroeconomic variables control for the relative aggressiveness of the bank over the business cycle. Of course, this beta model can also be used with variables for which the share model is also feasible. The automatic model selection algorithm discussed below will try both the share and beta models in those cases.

4. Peer analysis

In the previous section we discussed using industry-level data to forecast an individual bank's financial statements. However, when making forecasts or developing growth strategies, most banks do not compare themselves with the entire industry. Rather, as do firms in most industries, the typical bank compares itself with a small group of peers against whom it competes most aggressively for business. For a community or regional bank, its peers likely include banks located in the same geographic area. For a larger bank that operates nationally, its peer group may consist of other large banks that specialize in the same activities. In principle a bank might consider one group of banks as its peer with respect to some lines of business and another group of banks as its peer with respect to other lines. For the sake of discussion, however, we will assume the bank has a single peer group.

To compute the industry-level data, we aggregate the individual call reports from the more than 6,000 banks that constitute the industry. We can do the same aggregation for any combination of banks we want. Suppose we want to forecast the financial statement of bank A, and its primary competitors are banks B, C, D, E and F. We can create a set of financial statements for a hypothetical bank as if banks A, B, C, D, E and F were one bank.

Forecasting our peer group proceeds just as if we were forecasting a single bank, using the automatic model selection procedure we describe in section 6. We simply replace the bank-level data in that procedure with the data for our peer group. Doing that we obtain scenario-specific forecasts for all the variables in the CRF at the peer group level.

Forecasting bank A given the historical data and forecasts for its peer group also proceeds by using the algorithm described in section 6. Now we use the data for our bank A as the bank-level data, and we use the historical and forecast peer-group data in place of the industry data. In this sense, forecasting an individual bank based on industry data is in fact a special case of peer-group analysis where the entire industry is the peer group.

We could also make forecasts for banks B, C, D, E and F in the same way, allowing bank A to not only see how it and its entire peer group is likely to perform in various economic scenarios but also how each of its competitors will perform individually. We are doing nothing nefarious: All of the bank data are publicly available.

The market share approach is particularly potent here. Suppose bank A's managers think its residential mortgage assets will grow 20% over the next two years. In contrast, say our peer analysis shows that the entire group's residential mortgage assets will grow by only 5%. Our peer forecast makes clear to bank A's managers that the only way to achieve its 20% growth target is by increasing market share. If a substantial increase is not feasible, then management will have to trim its internal forecast.

5. Automatic model selection

To forecast each variable for a bank, we need a model that ties the market share to macroeconomic variables or that ties the bank-level variable to the industry aggregate and macroeconomic variables. Our algorithm fits dozens of candidate models and picks the model that provides the highest out-of-sample forecast accuracy. With more than 200 variables in the CRF, we fit at least 5,000 individual models when forecasting a bank or a peer group.

Macroeconomic factors

Whether we fit a share model or a beta model, we must control for macroeconomic factors that affect the variable being forecast. The typical modeler at a bank will search through dozens if not hundreds of macroeconomic variables, searching for just the right combination of regressors to explain the variable being forecast. Macroeconomic variables are highly correlated, and the multiple-testing problem pervades that type of modeling.

We instead use principal components analysis (PCA) to extract indexes of economic activity and use the top three indexes in our automatic model selection algorithm.

PCA is a data-reduction algorithm that extracts from a large set of input variables a set of principal components, variables that are by construction independent of one another but that jointly explain all the variation in the original set of variables. Typically the first few PCs can explain the bulk of the variation in the input data. Thus, PCA allows us to take a large set of macroeconomic variables and develop a few index variables that convey most of the information in that large set of (correlated) macroeconomic variables, eliminating the futile chore of finding just the right macroeconomic variables to forecast each line item on the bank's financial statements.

Table 1 shows the macroeconomic variables included in our PCA. Obviously nonstationary variables are converted to percentage changes, and we use data from 1987Q1 through 2016Q2. Table 2 shows that the first PCA explains nearly 66% of the variation in our macroeconomic variables, and the first three PCs jointly explain nearly 82%. The scree plot flattens after the third PC, so our automatic model selection algorithm uses just the first three.

A common criticism of PCA is that the resulting PCs are difficult to interpret and bear little resemblance to known variables. Charts 4, 5 and 6 show that is not a problem here. The first PC captures the secular decline and cyclical gyrations in interest rates over the past three decades; the PC forecasts in the three CCAR scenarios are similar to the specified paths for interest rates.

Table 1: Macroeconomic Variables Used in PCA

Variable	Functional form
CPI: Urban consumer - All items	% Change
Futures price: NYMEX light sweet crude oil - contract 1	Levels
Employment: Total nonagricultural	% Change
Gross domestic product	% Change
New-home sales: Single-family homes	% Change
Median existing single-family home price	% Change
Existing-home sales: Single-family	% Change
Unemployment rate	Levels
Retail sales: Retail sales and food services	% Change
New-vehicle sales: Cars and light trucks	Levels
S&P stock price index: 500 composite - monthly avg	% Change
Income: Disposable personal	% Change
Income: Personal - total	% Change
Moody's bond yield - Aaa corporate - 20+ yr maturities	Levels
Moody's bond yield - Aa corporate - 20+ yr maturities	Levels
Moody's bond yield - A corporate - 20+ yr maturities	Levels
Moody's bond yield - Baa corporate - 20+ yr maturities	Levels
CDs secondary market - 1 mo	Levels
CDs secondary market - 3 mo	Levels
CDs secondary market - 6 mo	Levels
Eurodollar deposits (London) - 1 mo	Levels
Eurodollar deposits (London) - 3 mo	Levels
Eurodollar deposits (London) - 6 mo	Levels
Federal funds rate	Levels
Three-mo Treasury bill yield - secondary market	Levels
Six-mo Treasury bill yield - secondary market	Levels
One-yr constant maturity Treasury security yield	Levels
Two-yr constant maturity Treasury security yield	Levels
Three-yr constant maturity Treasury security yield	Levels
Five-yr constant maturity Treasury security yield	Levels
Seven-yr constant maturity Treasury security yield	Levels
10-yr constant maturity Treasury security yield	Levels
30-yr constant maturity Treasury security yield	Levels
One-mo USD LIBOR deposit rate	Levels
Three-mo USD LIBOR deposit rate	Levels
Six-mo USD LIBOR deposit rate	Levels
12-mo USD LIBOR deposit rate	Levels
Bank prime rate	Levels

Source: Moody's Analytics

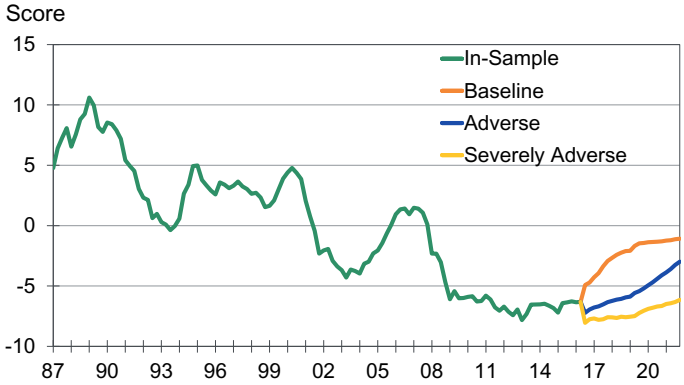
Table 2: Macroeconomic Variable PCA Results

Number of observations = 116

Component	Eigenvalue	Difference	Variance proportion	Cumulative variance proportion
1	25.060	21.361	0.660	0.660
2	3.699	1.446	0.097	0.757
3	2.252	0.574	0.059	0.816
4	1.678	0.510	0.044	0.860
5	1.168	0.258	0.031	0.891

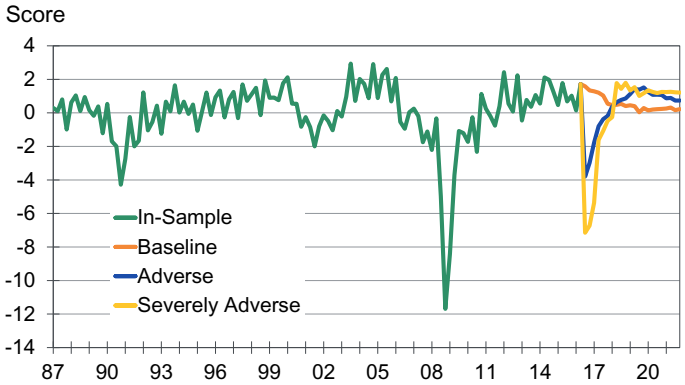
Source: Moody's Analytics

Chart 4: First Principal Component



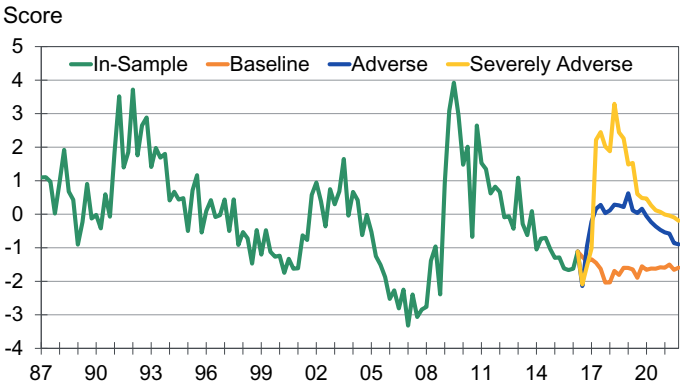
Source: Moody's Analytics

Chart 5: Second Principal Component



Source: Moody's Analytics

Chart 6: Third Principal Component



Source: Moody's Analytics

The second and third PCs reflect the overall business cycle. The second one applies the most weight to GDP, employment and income growth as well as vehicle sales, house prices, and the stock market. The third one is most sensitive to the unemployment rate, new- and existing-home sales, and vehicle sales; the third PC also picks up variation in long-term interest rates not explained by the first PC.

Model algorithm

We let b_t represent the bank-level variable being forecast and p_t be the industry-level analogue. t indexes calendar quarters. The three principal components are $P1_t$, $P2_t$, and $P3_t$. The bank's market share s_t equals b_t/p_t.

We perform preliminary screens on the bank-level variable before invoking the model-selection algorithm. In these cases the variable is not amenable to regression modeling. We set the forecasts equal to zero if five or fewer observations in history are nonzero or if five or fewer observations since 2005Q1 are nonzero. We set the forecasts equal to the historical mean if the coefficient of variation of the variable since 2005Q1 is less than 2%. Lastly, if there are fewer than 16 observations we set the forecasts to the mean of the most recent five observations.

In section 3 we discussed the share and beta models. The share model works only for variables that are positive, while the beta model can be used with all variables. A shortcoming of the beta model, though, is that if we use it to forecast a variable that must be nonnegative or positive there is no guarantee that the forecasts will observe that constraint. Our algorithm therefore encompasses four models:

- Share model for variables that must be positive;
- Beta model for unconstrained variables;
- Nonnegative beta model for variables that must be zero or greater; and
- Positive beta model for variables that must be positive.

For variables that are strictly positive, we consider share and positive beta models; for variables that are zero or positive, we use the nonnegative beta model; and for all other variables we use the general beta model.

Our algorithm rejects any candidate model for which the variance-covariance matrix of the parameter estimates is not of full rank. Although the three PCs are by construction mutually orthogonal, once we include lags of those variables and perhaps the industry-level variable, the regression can fail, particularly when we are modeling a variable with a short history.

Share model

For the share model we must have s_t between 0 and 1 in all forecast periods and scenarios. To ensure that, we model the logit of s_t, $\text{logit}(s_t) = \exp(s_t) / \{1 + \exp(s_t)\}$. We fit ordinary least squares regressions of $\text{logit}(s_t)$ on a constant and

- $P1_t$ and from zero to five lags
- $P2_t$ and from zero to five lags
- $P3_t$ and from zero to five lags
- $\text{logit}(s_{t-1})$

To reduce the number of regressions we must run, we use the same number of lags for each of the PC variables in the candidate models, and if we fit a model with k lags, we also include lags $k-1$, $k-2$, … We also fit regressions without the third PC or its lags, and we fit regressions without the lagged share terms.

Thus, for each variable to which the share model applies, we fit 6 x 2 x 2 = 24 candidate models: Zero through five lags the PCs, with and without the terms corresponding to $P3_t$,

and with and without logit(s_{t-1}). We discuss selecting the best model later.

The OLS regression will forecast logit(s_t) under the economic scenarios we specify, and we take the inverse logit to obtain the forecast share, ignoring the retransformation problem. Simple multiplication of the forecast share by the industry forecast gives the bank-level forecast.

The generalized linear model is often used to fit share models[5], but we instead use OLS regression because it is much faster. We use out-of-sample forecast evaluation to choose the best model, and our focus is on forecast accuracy, not causal inference. Therefore the statistical rigor of the GLM approach is unwarranted here.

Beta model

For the beta model, the forecasts are unrestricted and we fit OLS regressions of b_t on a constant and

- Industry variable p_t
- $P1_t$ and from zero to five lags
- $P2_t$ and from zero to five lags
- $P3_t$ and from zero to five lags
- b_{t-1}

As with the share model, we restrict the PC variables to have the same number of lags, and we include all lower-order lags. We also fit variants of the model where b_t, p_t, and b_{t-1} are first differenced; and we fit variants that omit $P3_t$ and its lags and that omit b_{t-1}.

For each variable there is a total of 6 x 2 x 2 x 2 = 48 candidate models: zero through five lags of the PCs, with and without the terms corresponding to $P3_t$, with and without b_{t-1}, and with b_t, p_t, and b_{t-1} specified in levels or in first differences. If a variable has fewer than 50 observations in the estimation sample, we do not consider the models specified in first differences, reducing the number of candidates to 24. We found that with relatively few observations those models often produced implausible forecasts.

Nonnegative beta model

For variables that are zero or positive we use the nonnegative model. Here the bank-level variables are typically positive in the estimation sample but are on rare occasions zero. Let c_t equal b_t with all zero values replaced with 0.001 times the smallest nonzero value of b_t in the estimation sample. We fit OLS regressions of log(c_t) on a constant and

- log(p_t)
- $P1_t$ and from zero to five lags
- $P2_t$ and from zero to five lags
- $P3_t$ and from zero to five lags

As with the share model we restrict the PC variables to have the same number of lags, and we include all lower-order lags. We do not fit models that include a lag of the regressand, nor do we fit models that first difference the bank- and industry-level variables. In preliminary testing we found that those versions of this model invariably performed poorly. Therefore, this model only has 6 x 2 = 12 variants: zero through five lags of the PCs and whether to include the terms corresponding to $P3_t$. To obtain forecasts we simply exponentiate the predicted values from the logarithmic regression, again ignoring the retransformation problem. The predicted values will never be zero though they can be arbitrarily small.

Positive beta model

For variables that are strictly positive we fit regression of log(b_t) on a constant and

- log(p_t)
- $P1_t$ and from zero to five lags
- $P2_t$ and from zero to five lags
- $P3_t$ and from zero to five lags
- log(b_{t-1})

As with the other models we restrict the PC variables to have the same number of lags, and we include all lower-order lags. Here we consider models with and without the lagged regressand though again we do not consider models in first differences because of their frequent poor performance. This model has 6 x 2 x 2 = 24 variants: zero through five lags of the PCs, with and without terms corresponding to $P3_t$, and with and without log(b_{t-1}).

Choosing the best model

For each of the more than 200 variables we forecast, we have as many as 48 candidate models from which to choose. For most variables we use an out-of-sample back-test procedure. Suppose a variable b_t has observations for the entire history, from 1992Q4 through 2016Q2. We use observations from 1992Q4 through 2014Q1 and fit the appropriate share or beta models. For any candidate model, let f_t represent the forecast from that model. We compute the mean squared forecast error as

$$\text{MSE} = \sum (f_t - b_t)^2 / 9$$

where the summation runs from 2014Q2 through 2016Q2. We then choose the model with the lowest MSE. To obtain our final forecasts under all economic scenarios we refit the model using all data from 1992Q4 through 2016Q2. We chose a nine-quarter evaluation window to match the nine-quarter forecast horizon of CCAR stress tests.

Unfortunately not every variable in the CRF has history extending back to 1992Q4. For variables that begin in 2010Q1 or earlier, we use observations through 2014Q1 to fit the candidate models and perform out-of-sample validation using the last nine quarters' data. Fitting models with as few as 16 observations is not ideal, but our preliminary screens and rejection of any model with a rank-deficient VCV matrix help guard against egregious forecasts. For these variables, the most pertinent risk is overfitting: The chosen candidate models may perform well in the out-of-sample model selection, but when those models are refit using data through 2016Q2 the parameter estimates could change substantially and render the final scenario forecasts inaccurate. Nevertheless, the CRF forecasts of these variables do provide managers a starting point for qualitative forecasts or management overlays.

For variables that start after 2010Q1, we use the first 16 observations to fit the candidate models. We still select the final model based on MSEs computed over the 2014Q2 through 2016Q2 window, but these are no longer true out-of-sample statistics, as some observations are used both for model fitting and forecast evaluation.

5 L. E. Papke and J. M. Wooldridge, "Econometric methods for fractional response variables with an application to 401(k) plan participation rates," *Journal of Econometrics, 145*, (1996): 121-33.

6. Benchmarking

We performed a benchmark exercise to determine the accuracy of our automatic model selection algorithm. The results of this exercise will also be used to determine whether proposed changes to the algorithm should be accepted. We evaluate forecast accuracy in terms of the nine-quarter window spanning 2014Q2 through 2016Q2. In order to have sufficient data to fit candidate models and choose the optimal model based on the out-of-sample forecast evaluation described in the previous section, we focus on variables that are available for the entire 1992Q4-2016Q2 span of call report data. Table 3 describes the allocation of data in detail.

We used 250 banks in our benchmarking exercises, including 31 of the banks subject to the 2016 CCAR tests and 61 banks subject to the 2016 DFAST tests. We excluded two CCAR banks and 11 DFAST banks because their historical call report data did not extend back to 1992Q4. Our sample also includes random selections of 100 banks with between $1 billion and $10 billion in assets as of 2016Q2 and 58 banks with assets under $1 billion. By sampling across the size spectrum, we can assess how well our algorithm works on different types of banks.

We focused our attention on variables that are generally available throughout the entire history of the call reports and were part of the CRF library as of August 2016. Because this exercise is focused on evaluating the regression modeling algorithm, we also ignored variables that are often zero or unreported for some banks and hence likely to be flagged by our preliminary screens. We thus focus on 51 variables here.

Table 3: Benchmark Analysis Data Allocation

Span	Observations	Use
1992Q4 - 2011Q4	77	Candidate model estimation
2012Q1 - 2014Q1	9	Model selection
1992Q4 - 2014Q1	86	Final model estimation
2014Q2 - 2016Q2	9	Forecast evaluation

Source: Moody's Analytics

We use an absolute loss criterion to evaluate performance. The modeling algorithm selects the optimal model based on a squared loss criterion, so using a squared loss criterion in our benchmark analysis might overstate the performance of our algorithm. In particular, for each bank and for each variable we compute the symmetric mean absolute percentage error, or sMAPE, of the forecast over the 2014Q2-2016Q2 nine-quarter window, where sMAPE is defined as

$$\text{sMAPE} = 100 \cdot 2/9 \cdot \sum |f_t - b_t| / (|f_t| + |b_t|)$$

and where $|x|$ denotes the absolute value of x.

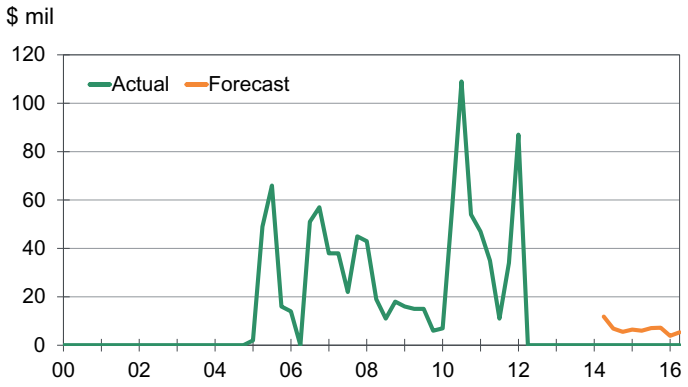
A shortcoming of the sMAPE measure is that if either the actual or forecast values are equal to zero for the entire evaluation window but the other values are not, then it will equal its theoretical maximum of 200%. Chart 7 shows an example. Here the model forecasts positive values, which is sensible, given that from 2005Q1 through 2012Q1 the actual value was zero only once. However, the actual values are zero throughout the evaluation window from 2014Q2 through 2016Q2. In other cases, where the series is often zero, the model can produce forecasts that are (essentially) zero while the actual values are positive in the evaluation window; these cases also result in sMAPE values that are (essentially) 200%.[6] To reiterate, obtaining an sMAPE of 200% represents a shortcoming of the sMAPE metric, not the forecast model. We therefore ignore results for which either all nine actual values or all nine forecast values equal zero.

Table 4 shows the distribution of sMAPE values across banks for the variables in focus. The automatic model selection algorithm works well for most variables, with median sMAPEs well under 10%. As one would expect, variables that tend to be more stable are forecast more accurately.

Some of the sMAPE statistics stand out, but closer inspections show the forecasts were reasonable, given the information up to the date at which they were made. For example, Chart 8 shows the forecast for noncurrent loans and

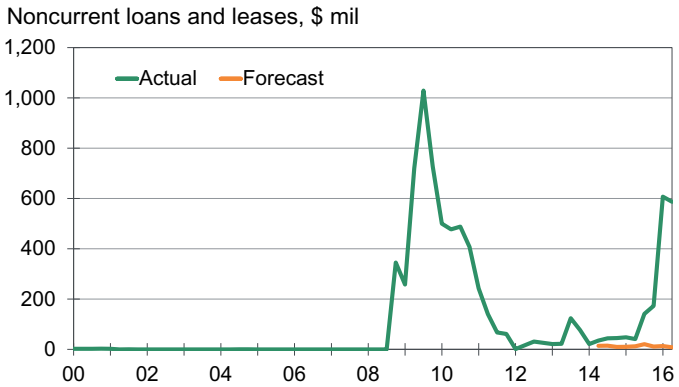
6 We say "essentially" because the positive beta model will never produce a forecast that is identically zero. When we ignore results for which the forecasts are zero, we in fact use the criterion that the mean of the forecasts is less than 0.05% of the historical mean of the series.

Chart 7: Zeroes Create Issues for sMAPE



Sources: FDIC Statistics on Depository Institutions, Moody's Analytics

Chart 8: Reasonable Forecast but High sMAPE



Sources: FDIC Statistics on Depository Institutions, Moody's Analytics

Table 4: sMAPE Results of Benchmarking Exercise

Variable	Number of banks	Percentile 10	25	50	75	90
Avg assets, quarterly	250	1.64	2.75	5.07	10.12	21.87
Avg assets	250	1.78	2.80	5.21	11.15	18.64
Total assets	250	1.87	3.10	5.72	9.96	17.80
Long-term assets (5+ yrs)	247	4.23	8.12	14.65	28.99	49.27
Adjusted avg assets for leverage capital purposes	247	1.68	2.85	4.96	9.40	18.53
Bank premises and fixed assets	245	2.42	4.15	7.88	15.55	42.80
Cash & balances due from depository institutions	250	18.04	26.01	38.91	57.08	82.96
Core deposits	247	2.35	3.67	6.30	10.55	18.21
Demand deposits	245	7.35	10.44	20.48	40.49	67.41
Total deposits	250	2.02	3.36	6.27	10.59	19.39
Deposits held in domestic offices	250	2.11	3.45	6.44	11.15	19.08
Interest-bearing deposits	247	2.29	3.65	6.85	13.51	23.68
Interest-bearing deposits in domestic offices	247	2.33	3.90	7.42	14.42	24.25
Estimated insured deposits	250	2.21	3.93	7.39	13.04	23.73
Noninterest-bearing deposits	246	4.14	5.87	10.38	20.18	37.84
Avg equity	247	1.25	2.78	5.73	12.33	21.54
Bank equity capital	247	1.62	3.14	6.32	12.70	20.75
Common stock	223	0.00	0.00	4.29	26.16	94.78
Surplus	235	0.23	2.15	7.14	17.32	42.80
Total equity capital	247	1.54	2.97	6.45	12.77	20.63
Undivided profits	245	2.60	5.57	12.86	38.46	92.30
Earning assets	247	1.91	3.18	6.08	10.48	19.51
All other assets	250	4.04	6.61	13.09	27.50	48.03
All other liabilities	250	9.01	13.49	23.06	38.03	63.44
Individuals, partnerships and corporations	250	2.06	3.49	6.60	12.31	23.04
IRAs and Keogh plan accounts	242	2.39	4.43	8.37	16.85	30.31
Total liabilities	250	2.09	3.48	6.36	11.35	23.00
Total liabilities and capital	250	1.87	3.10	5.72	9.96	17.80
Loan loss allowance	247	4.25	7.04	13.41	24.95	42.56
Loans to individuals	243	4.36	7.92	15.75	37.34	89.40
Total loans and leases	250	2.37	3.91	7.29	14.49	32.91
Net loans and leases	250	2.21	3.93	7.22	14.68	32.74
Noncurrent loans and leases	249	13.93	23.97	41.64	67.20	107.75
Nontransaction accounts	250	2.33	3.73	6.67	13.57	31.13
Time deposits less than $100,000 - amount ($)	242	2.92	6.31	12.23	26.14	52.44
Nontransaction accounts of individuals, partnerships and corporations	250	2.37	3.98	7.05	14.12	29.05
Money market deposit accounts	240	3.89	5.75	12.06	21.54	43.13
Other savings deposits (ex money market deposit accounts)	241	3.32	6.05	12.85	26.57	51.38
Total time deposits	245	4.77	7.70	12.64	25.91	46.58
Time deposits of $100,000 or more - amount ($)	248	7.68	12.08	20.37	34.46	68.60
Total employees (full-time equivalent)	246	1.94	2.98	5.89	12.55	24.40
Income earned, not collected on loans	247	3.61	6.45	12.27	20.54	43.16
Tier one (core) capital	247	1.34	2.64	5.57	11.80	25.03
Tier two risk-based capital	247	5.52	8.31	16.27	27.76	51.11
Total securities	247	4.01	7.52	14.94	25.92	49.89
Transaction accounts	249	6.06	9.52	15.34	31.40	57.98
Transaction accounts of individuals, partnerships and corporations	249	5.49	10.63	17.73	38.68	66.91
Total time and savings deposits	247	3.92	6.80	9.35	15.62	27.21
Total unused commitments	249	4.36	7.97	15.09	29.78	56.44
Unused loan commitments	249	4.23	7.51	14.53	30.50	55.63
Volatile liabilities	245	10.23	17.82	31.39	54.24	102.17

Source: Moody's Analytics

leases for the bank with the largest sMAPE for that variable. The variable had been nil up until the Great Recession, spiked higher, and then came back down. Given the mostly benign economic conditions in 2014 and 2015, the forecast is therefore reasonable. Nevertheless, the bank reported sharply higher noncurrent loans in 2015 and 2016.

Table 5 shows the median sMAPE for each variable, disaggregated by type of bank. In general our model selection algorithm performs equally well regardless of the size of the bank. For 20 of the 51 variables, the median sMAPE is lower for CCAR banks than for the small banks with less than $1 billion in assets. However, for many of those variables such as the number of employees or net loans and leases the differences are trivial. Variables such as unused loan commitments are arguably more relevant for larger banks, and the noisy data for smaller banks is inherently difficult to forecast.

7. Conclusion

Most businesses forecast sales and profitability by considering industry-level trends and their position within the industry. The banking industry has been an exception, but the Moody's Analytics Bank Call Report Forecasts database now provides banks with the industry forecasts needed to apply that approach.

Our methodology is built from publicly available bank-level data extending back to 1992Q4. Together with our merger adjustment algorithm, that provides us with clean, consistent data with a much longer history than most banks' internal data warehouses provide. Moreover, that allows us to create custom peer groups and make forecasts not just for some banks but for their competitors as well.

We use an automatic model selection procedure to produce bank-level forecasts under various economic scenarios. Benchmark analysis shows that our procedure produces accurate forecasts regardless of the size of the bank or peer group being modeled.

Table 5: sMAPE Results by Type of Bank

	Type of bank			
Variable	**CCAR**	**DFAST**	**$1 bil-$10 bil**	**< $1 bil**
Avg assets, quarterly	5.31 (31)	5.05 (61)	5.32 (100)	4.83 (58)
Avg assets	5.09 (31)	7.50 (61)	4.85 (100)	4.36 (58)
Total assets	6.96 (31)	6.15 (61)	5.66 (100)	4.48 (58)
Long-term assets (5+ yrs)	12.92 (31)	14.14 (59)	14.65 (99)	16.04 (58)
Adjusted avg assets for leverage capital purposes	4.18 (31)	4.98 (59)	5.57 (99)	4.57 (58)
Bank premises and fixed assets	5.78 (30)	7.85 (58)	8.03 (99)	8.00 (58)
Cash and balances due from depository institutions	26.52 (31)	40.30 (61)	43.47 (100)	37.58 (58)
Core deposits	5.05 (31)	7.72 (59)	6.34 (99)	5.93 (58)
Demand deposits	20.91 (31)	28.06 (58)	20.81 (98)	15.44 (58)
Total deposits	6.14 (31)	7.47 (61)	6.81 (100)	5.16 (58)
Deposits held in domestic offices	6.20 (31)	8.00 (61)	6.83 (100)	5.38 (58)
Interest-bearing deposits	6.51 (31)	8.38 (59)	7.65 (99)	5.63 (58)
Interest-bearing deposits in domestic offices	7.82 (31)	8.66 (59)	7.67 (99)	5.51 (58)
Estimated insured deposits	7.44 (31)	9.24 (61)	8.02 (100)	5.44 (58)
Noninterest-bearing deposits	7.32 (31)	10.56 (58)	10.06 (99)	13.29 (58)
Avg equity	4.66 (31)	5.93 (59)	6.91 (99)	3.46 (58)
Bank equity capital	4.37 (31)	7.36 (59)	7.74 (99)	4.23 (58)
Common stock	10.35 (30)	8.17 (54)	6.54 (91)	0.00 (48)
Surplus	5.98 (31)	9.65 (59)	8.69 (96)	3.35 (49)
Total equity capital	3.58 (31)	8.15 (59)	7.83 (99)	4.09 (58)
Undivided profits	11.09 (30)	15.51 (59)	16.22 (99)	9.52 (57)
Earning assets	6.37 (31)	7.48 (59)	6.08 (99)	4.55 (58)
All other assets	9.21 (31)	13.87 (61)	12.27 (100)	15.86 (58)
All other liabilities	17.88 (31)	22.85 (61)	23.06 (100)	24.94 (58)
Individuals, partnerships and corporations	5.48 (31)	6.70 (61)	8.09 (100)	4.62 (58)
IRAs and Keogh plan accounts	8.50 (30)	8.07 (57)	8.59 (99)	8.37 (56)
Total liabilities	6.90 (31)	7.67 (61)	6.83 (100)	4.93 (58)
Total liabilities and capital	6.96 (31)	6.15 (61)	5.66 (100)	4.48 (58)
Loan-loss allowance	14.94 (31)	15.15 (59)	13.74 (99)	11.35 (58)
Loans to individuals	12.32 (31)	15.58 (58)	17.84 (98)	14.93 (56)
Total loans and leases	8.64 (31)	7.38 (61)	7.60 (100)	5.58 (58)
Net loans and leases	5.06 (31)	7.57 (61)	7.95 (100)	5.55 (58)
Noncurrent loans and leases	28.29 (31)	37.51 (61)	41.14 (99)	74.18 (58)
Nontransaction accounts	5.77 (31)	9.12 (61)	6.45 (100)	5.51 (58)

Table 5: sMAPE Results by Type of Bank (Cont.)

	Type of bank			
Variable	**CCAR**	**DFAST**	**$1 bil-$10 bil**	**< $1 bil**
Time deposits less than $100,000 - amount ($)	33.48 (29)	17.81 (57)	12.92 (99)	7.88 (57)
Nontransaction accounts of individuals, partnerships and corporations	5.97 (31)	8.70 (61)	7.50 (100)	5.97 (58)
Money market deposit accounts	7.03 (31)	12.94 (58)	11.68 (99)	14.31 (52)
Other savings deposits (ex money market deposit accounts)	14.65 (29)	13.42 (57)	13.15 (98)	10.78 (57)
Total time deposits	26.54 (31)	13.30 (58)	13.05 (99)	9.35 (57)
Time deposits of $100,000 or more - amount ($)	35.36 (31)	21.73 (60)	18.76 (100)	16.13 (57)
Total employees (full-time equivalent)	3.84 (31)	6.10 (58)	7.48 (99)	4.71 (58)
Income earned, not collected on loans	9.48 (31)	13.74 (59)	11.52 (99)	13.81 (58)
Tier one (core) capital	5.10 (31)	7.54 (59)	6.02 (99)	3.69 (58)
Tier two risk-based capital	22.53 (31)	15.12 (59)	17.41 (99)	14.02 (58)
Total securities	10.54 (30)	15.53 (59)	14.21 (100)	16.29 (58)
Transaction accounts	14.61 (31)	23.78 (60)	20.04 (100)	10.54 (58)
Transaction accounts of individuals, partnerships and corporations	17.32 (31)	29.03 (60)	19.78 (100)	12.50 (58)
Total time and savings deposits	8.75 (31)	10.43 (59)	9.84 (99)	8.68 (58)
Total unused commitments	7.27 (31)	14.58 (61)	14.16 (100)	27.20 (57)
Unused loan commitments	7.21 (31)	13.06 (61)	15.68 (100)	27.44 (57)
Volatile liabilities	23.91 (31)	26.57 (59)	38.09 (99)	37.85 (56)

(Number of observations in parentheses)

Source: Moody's Analytics

About the Author

Brian Poi is a director in the Specialized Modeling Group at Moody's Analytics in West Chester PA, where he develops new products for forecasting and stress testing purposes, leads external model validation projects, and supervises econometric model development for the Moody's Analytics U.S. economic forecast model. He also provides thought leadership and guidance on the use of advanced statistical and econometric methods in economic forecasting applications. In his prior role he developed a variety of credit loss, credit origination and deposit account models for use in both strategic planning and CCAR/DFAST environments. Before joining Moody's Analytics, Dr. Poi was an econometric developer and director of professional services at StataCorp LP, a leading provider of statistical analysis software. He received his PhD and MA in economics from the University of Michigan after graduating magna cum laude from Indiana University.

About Moody's Analytics

Moody's Analytics helps capital markets and credit risk management professionals worldwide respond to an evolving marketplace with confidence. With its team of economists, the company offers unique tools and best practices for measuring and managing risk through expertise and experience in credit analysis, economic research, and financial risk management. By offering leading-edge software and advisory services, as well as the proprietary credit research produced by Moody's Investors Service, Moody's Analytics integrates and customizes its offerings to address specific business challenges.

Concise and timely economic research by Moody's Analytics supports firms and policymakers in strategic planning, product and sales forecasting, credit risk and sensitivity management, and investment research. Our economic research publications provide in-depth analysis of the global economy, including the U.S. and all of its state and metropolitan areas, all European countries and their subnational areas, Asia, and the Americas. We track and forecast economic growth and cover specialized topics such as labor markets, housing, consumer spending and credit, output and income, mortgage activity, demographics, central bank behavior, and prices. We also provide real-time monitoring of macroeconomic indicators and analysis on timely topics such as monetary policy and sovereign risk. Our clients include multinational corporations, governments at all levels, central banks, financial regulators, retailers, mutual funds, financial institutions, utilities, residential and commercial real estate firms, insurance companies, and professional investors.

Moody's Analytics added the economic forecasting firm Economy.com to its portfolio in 2005. This unit is based in West Chester PA, a suburb of Philadelphia, with offices in London, Prague and Sydney. More information is available at www.economy.com.

Moody's Analytics is a subsidiary of Moody's Corporation (NYSE: MCO). Further information is available at www.moodysanalytics.com.

About Moody's Corporation

Moody's is an essential component of the global capital markets, providing credit ratings, research, tools and analysis that contribute to transparent and integrated financial markets. **Moody's Corporation** (NYSE: MCO) is the parent company of Moody's Investors Service, which provides credit ratings and research covering debt instruments and securities, and **Moody's Analytics**, which encompasses the growing array of Moody's nonratings businesses, including risk management software for financial institutions, quantitative credit analysis tools, economic research and data services, data and analytical tools for the structured finance market, and training and other professional services. The corporation, which reported revenue of $3.5 billion in 2015, employs approximately 10,400 people worldwide and maintains a presence in 36 countries.