Federal Reserve Bank of Chicago

BankCaR (Bank Capital-at-Risk): A credit risk model for US commercial bank charge-offs

Jon Frye and Eduard Pelz

WP 2008-03

BankCaR (Bank Capital-at-Risk): A credit risk model for US commercial bank charge-offs

Jon Frye
Senior Economist
Federal Reserve Bank of Chicago
230 South LaSalle Street
Chicago, IL 60604
Jon.Frye@chi.frb.org
312-322-5035

Eduard Pelz
Senior Supervision Analyst
Federal Reserve Bank of Chicago
230 South LaSalle Street
Chicago, IL 60604
Eduard.Pelz@chi.frb.org
312-322-6018

April 22, 2008

The authors are indebted to numerous individuals at the Federal Reserve who have given contributions and insightful comments. These include members of the Research and Bank Supervision Departments of the Federal Reserve Banks of Chicago, Richmond, and Dallas, and the Federal Reserve Board of Governors. The views expressed are the authors' and do not necessarily represent the views of the management of the Federal Reserve Bank of Chicago or the Federal Reserve System.

BankCaR (Bank Capital-at-Risk):
A credit risk model for US commercial bank charge-offs

Abstract

BankCaR is a credit risk model that forecasts the distribution of a commercial bank's charge-offs. The distribution depends only on systematic factors; BankCaR takes each bank and projects its expected charge-off across a distribution of good years and bad years. Since most bank failures occur in bad years, this analysis has promise for both banks and bank supervisors.

In BankCaR, charge-offs depend on the bank's loan balances and the charge-off rates of twelve categories of lending. A joint distribution of the twelve charge-off rates is calibrated to a long history of regulatory reporting data. Applied to the US banking system, BankCaR finds that credit risk is rising and is concentrated most significantly in construction lending. Applied to individual banks, BankCaR efficiently identifies those that have an adverse combination of credit risk and capital.

BankCaR uses publicly available regulatory reporting data, the most common credit portfolio model, and standard quantitative techniques. These generic qualities can provide a standard of comparison between banks. They also can provide an individual commercial bank with a benchmark for more elaborate vended credit models.

Keywords: credit risk, risk screening, loan charge-offs, validation

JEL classification: G32, G21, G38

Introduction and summary

BankCaR (Bank Capital-at-Risk) is a credit risk model that forecasts the distribution of charge-offs at commercial banks in the United States. Its goal is to improve the identification, quantification, and management of inherent bank credit risk. BankCaR can supplement a bank's internal credit risk model or benchmark credit risk in the absence of other models.

BankCaR forecasts risks rather than outcomes. Traditionally, the analysis of bank risk involves a forecast of what is ***likely*** to happen to a bank, a sector, or the macro economy. Instead, BankCaR forecasts what ***can*** happen under a broad range of circumstances. BankCaR's focus on risk represents a departure from traditional analyses of bank losses.

Applied to the US banking system, BankCaR finds that credit risk is rising, and that it is concentrated most significantly in construction lending. Applied to individual banks, BankCaR identifies those that have an adverse combination of credit risk and capital.

The paper begins with a brief discussion of the data employed. The data are gathered from regulatory filings that are available on the public record since 1984. These filings do not contain all the information that is available to bank managers and bank examiners. Where detailed information is not available, BankCaR assumes uniformity between banks and over time.

An overview describes the model in general terms. BankCaR extends the model employed in Basel II to include twelve risk factors corresponding to twelve categories of lending. The probability distribution of a bank's annual charge-off depends on the bank's twelve loan balances and on the statistical distribution of the twelve charge-off rates. The assumptions are discussed in detail. Most of the assumptions arise from the nature of the data in the regulatory filings. Following the overview, the mathematics of the model and its calibration are presented.

Two sections report results at year-end 2006. First, BankCaR is applied to the aggregated US banking system. The greatest potential for loss stems from three categories: consumer, commercial and industrial (C&I), and construction. Construction becomes increasingly likely to dominate at greater levels of overall loss. This result departs from historical data, in which construction charge-offs have never played the dominant role. Thus, BankCaR predicts that a year with unusually high charge-offs will have an unusual primary source, namely, charge-offs on construction loans.

Next, BankCaR is applied separately to each US commercial bank. The majority of banks, if exposed to stressful conditions, charge off more construction loans than any other category. Scenarios those that cause banks to become undercapitalized tend to produce high levels of construction charge-offs.

BankCaR efficiently identifies banks that have an adverse combination of credit risk and capital.

Bank loans and charge-offs

BankCaR begins its analysis with the outstanding loan balances in twelve categories. The categories appear in Table 1.

Table 1. Loan categories and Call Report data elements

Category	Call Report data elements
C&I	Commercial and industrial loans
Consumer	Credit cards Other revolving credit plans Other consumer loans (includes single payment, installment, and all student loans)
Other	Other loans Loans to foreign governments and official institutions (including foreign central banks) Obligations (other than securities and leases of states and political subdivisions in the US)
Depository Institutions	Loans to depository institutions and acceptances of other banks To commercial banks in the US To other depository institutions in the US To banks in foreign countries
Lease Financing	Lease financing receivables (net of unearned income)
Agriculture	Loans to finance agricultural production and other loans to farmers
Construction	Construction, land development, and other land loans
Non-farm non-residential	Secured by nonfarm nonresidential properties
Multifamily	Secured by multifamily (5 or more) residential properties
Farm	Secured by farmland (including farm residential and other improvements)
1-4 Revolving	Revolving, open-end loans secured by 1–4 family residential properties and extended under lines of credit
1-4 Other	Closed-end loans secured by 1–4 family residential properties
Tier 1	Tier 1 Capital
ALLL	Allowance for Loan and Lease Losses
Total Assets	Total Assets

Loans are taken from Schedule RC-C *Loans and Lease Financing Receivables*, charge-offs from RI-B *Charge-offs and Recoveries on Loans and Leases and Changes in Allowance for Loan and Lease Losses*, Tier 1 capital from RC-R *Regulatory Capital* and total assets and ALLL from RC *Balance Sheet*. See http://www.chicagofed.org/economic_research_and_data/commercial_bank_data.cfm.

The data source is the *Consolidated Reports of Condition and Income*, generally referred to as the call reports. Call reports are an authoritative source of both the current loan balances for a bank and the historical charge-off rates for the banking system. Banks are required by law to make timely and accurate call report filings, and bank examiners verify that the filings reflect a bank's accounts. Call report data are available to the public and useful beginning 1984—a period significantly longer than most other credit loss data sets.

As shown in Table 1, some categories are combinations of several call report data elements. Over time data elements have evolved, usually to provide greater detail. This is most notably the case for the six real estate categories that became separately available only beginning in 1991.[1]

BankCaR uses the category data in two distinct ways: a bank's outstanding dollar balances are the amounts subject to charge-off, and historical US aggregate annual gross charge-off rates are used to calibrate the probability model. The

[1] Real estate consists of four commercial real estate categories (construction, farm, multifamily, and nonfarm nonresidential) and two residential categories (one-to-four-family revolving residential and one-to-four-family other residential).

combination of a bank's current loans and the distribution of charge-off rates produce assessments of bank risk.

Figure 1. Total bank loans outstanding with proportions

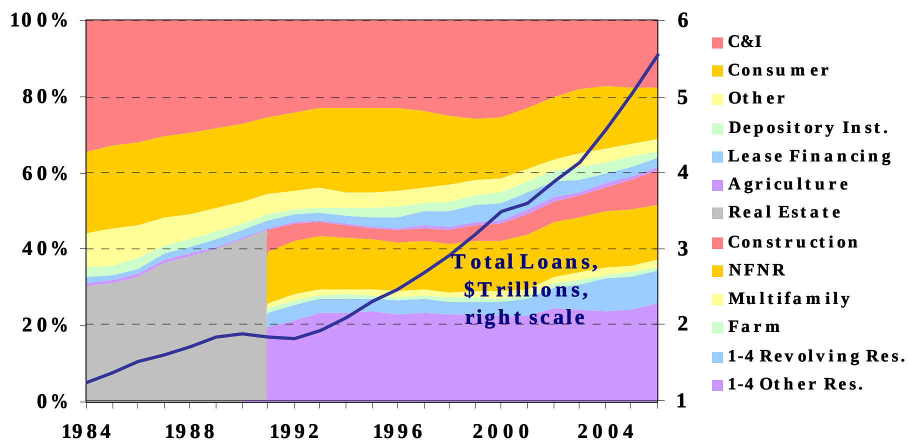


Figure 1 presents the recent history of outstanding loan balances for all US commercial banks. Loans grow 6.9% per year on average over the sample period. Also shown is the decomposition of loans into the twelve categories. Over the sample period, the proportion devoted to C&I declines by about half while proportion devoted to real estate approximately doubles.

Figure 2. Selected aggregate charge-off rates

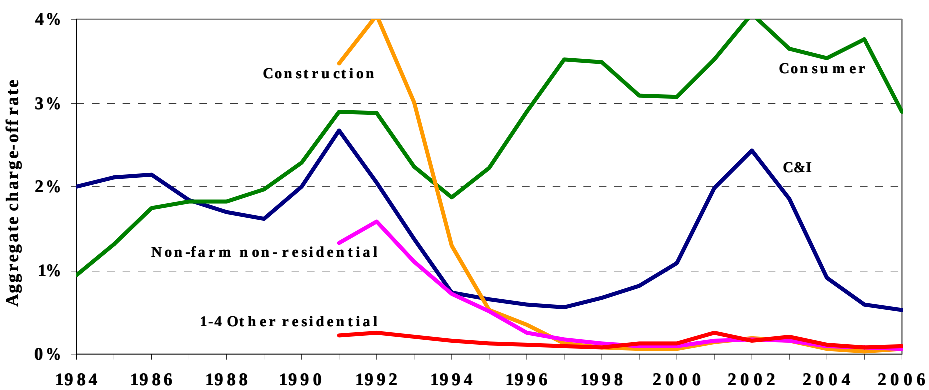


Figure 2 shows annual aggregate gross charge-off rates for US commercial banks in five selected categories. Several characteristics stand out. Different categories have different means, ranging from 0.15% (one-to-four-family other residential) to 2.68% (consumer). Dispersion around these means also differ from category to category. Further, the timing of cycles of different categories is not well synchronized, for example, the 2001 recession has little or no effect on real estate charge-offs; consumer lending has cycles of charge-off rates apparently unrelated to other categories.

Overview of BankCaR

This section begins with an intuitive motivation for the statistical model underlying BankCaR. It then describes BankCaR's distinguishing features, which are correlated factors and charge-off rate scenarios. The section concludes with definitions of the risk measures produced by BankCaR.

Figures 1 and 2 suggest the features that a statistical model of bank charge-off rates should possess. The model should analyze at least to the category level, because the categories have different average charge-off rates and the relative weights of the categories can change. Categories should be allowed to differ not only in mean but also in dispersion around the mean. Finally, categories should be allowed to vary together, but with less than perfect synchrony. A single factor model (such as employed in Basel II) cannot capture the sometimes contrary behaviors of charge-off rates in different categories.

The model underlying BankCaR is designed to match these broad features. To allow for lack of synchrony, BankCaR employs twelve separate, imperfectly correlated risk factors. Each factor controls a category charge-off rate. To allow different categories to have different means and different dispersions, each charge-off distribution is governed by two parameters.

A random draw of twelve possible charge-off rates is referred to as a charge-off scenario. Applied to a bank, a scenario produces an amount charged off. BankCaR employs 100,000 scenarios in total. The set of 100,000 possible charge-off amounts constitute the estimate of the bank's charge-off distribution.

Figure 3. From correlated normal factors to BankCaR scenarios

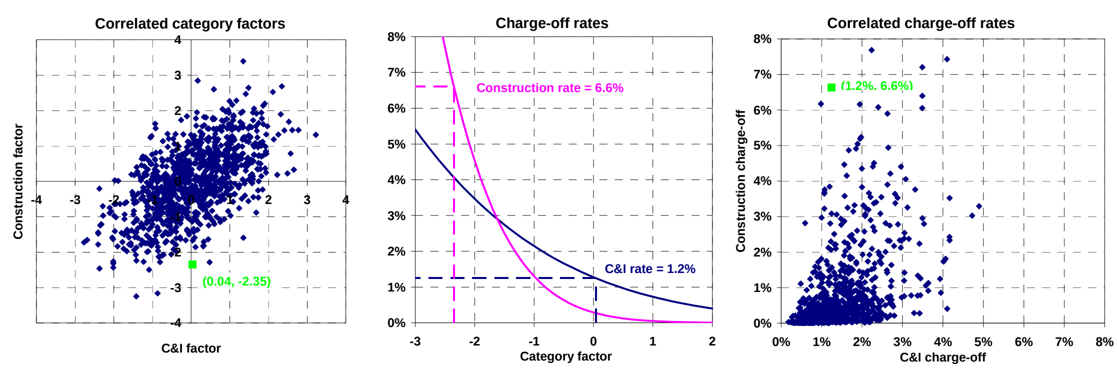


Figure 3 illustrates the production of 1,000 scenarios for two categories, C&I and construction. BankCaR uses the same procedure to produce 100,000 scenarios involving all twelve category charge-off rates.

The panel at the left shows 1,000 random draws from a pair of correlated standard normal factors. They exhibit the degree of correlation calibrated between C&I and construction, which is 59%. For illustrative purposes, a point is chosen and highlighted with factor values of -2.35 (construction) and 0.04 (C&I).

The two curves in the center panel convert these two factors into charge-off rates. There are of course twelve such curves within BankCaR, and all are calibrated to the data. The illustrative draw of the construction factor produces the sizable construction charge-off rate of 6.6%. The near-zero draw of the C&I factor produces a middling C&I charge-off rate of 1.2%. This pair of simulated charge-off rates is highlighted in the panel at the right.

The panel at the right shows the 1,000 scenarios that correspond to points in the panel at the left. According to the right panel, there will many years with near-zero construction charge-off rates and a few years with highly elevated rates; the charge-off rates for C&I are not forecast at either extreme. These stylized facts reflect the general patterns apparent in the historical rates shown in Figure 2.

Normally, the mean of each simulated factor is equal to zero, and the mean of each simulated charge-off rate is equal to the average historical charge-off rate. If the next year's charge-off rate is expected to differ from the historical average, a fixed amount can be added to draws of the associated factor. This adjustment allows for nontrivial forecasting while maintaining other aspects of the statistical distribution.

When new data become available, the statistical parameters are recalibrated and new scenarios are generated. The same set of scenarios is used for all analyses performed in a year. Using the same scenarios for different banks removes a source of randomness that would otherwise affect comparisons between banks.

BankCaR produces several measures and indicators of risk: a bank's Capital-at-Risk (CaR), its characteristic scenario, its risk type, and its stressed capital.

CaR is defined as the 99.5^{th} percentile of a bank's loss. Operationally, this equals the 500^{th} worst among the 100,000 possible losses, stated as a percent of the bank's total assets. CaR is the credit portfolio equivalent of Value-at-Risk, commonly used to the measure the risk of portfolios of traded financial instruments.

A bank's characteristic scenario portrays the average conditions under which loss equals CaR. Conceptually, the characteristic scenario is the mathematically expected scenario that produces loss equal to CaR. Operationally, BankCaR takes the average among a subset of its scenarios. Scenario losses are sorted and averages taken of the greatest two losses, the greatest three, and so forth. One of these averages is closest to CaR. The average of the associated scenarios is the characteristic scenario. In practice the characteristic scenario is the average of 1-2% of all scenarios. Thus, the characteristic scenario produces loss equal to CaR and serves as a portrayal of the average scenario that produces loss at that level.

Taken together, CaR and the characteristic scenario provide a summary of the bank's loss distribution. An even simpler summary is provided by CaR and risk type. A bank's risk type is the category that contributes the bank's greatest charge-off when the characteristic scenario is active. A given bank's risk type is not the whole story since two or more categories might contribute nearly equally to loss, but risk type can nonetheless be a useful summary.

BankCaR compares a bank's CaR to its financial resources. Specifically, stressed capital is defined as equal to tier 1 capital plus allowances for loan and lease losses (ALLL) less CaR.[2] Presumably, a bank having greater stressed capital would be better able to survive a year of high credit loss. Finally, BankCaR assigns a degree of inherent credit risk (High, Above Average, Average, or Low) based on the bank's stressed capital compared to other banks.

Assumptions and limitations

Like all statistical models, BankCaR makes numerous assumptions. The most important assumption is that every bank is vulnerable to the system-wide charge-off scenario, irrespective of the bank's location, size, or other characteristics. Although this assumption excludes some sources of bank risk, it facilitates a direct comparison between different institutions. Naturally, nothing reported by BankCaR can supplant detailed knowledge of an institution, its history, and its competitive environment.

Some assumptions follow from the nature of call report data. Call report category data are not distinguished by risk; therefore, BankCaR assumes that all loans within a category are uniform in obligor quality, outstanding amount, and potential for recovery. Call reports reflect only category-level outstanding loan balances. BankCaR does not model gains or losses from investment accounts or from contingent exposures such as loan commitments and off-balance sheet derivatives. Call reports do not contain estimates of the marked-to-market values of loan portfolios; BankCaR is calibrated only to amounts actually charged off.

Parameter values are assumed to be constant over time, as would be the case if underwriting standards were unchanged. In calibrating parameters, two choices are believed to have minor effects. The gross charge-off rate is used, in part because net charge-offs can be negative for technical reasons. The annual time step is used, in part because unobserved category factors are autocorrelated at the quarterly time step.

BankCaR does not quantify many variables that are important to overall bank risk, such as liquidity, control structures, technical and managerial quality, off-balance sheet exposures, hedging activities, and so forth. In particular, BankCaR

[2] Not all of ALLL might be available to absorb charge-offs, because some provisions in ALLL are associated with specific loans that might not default. The distinction between specific provisions (FAS-5) and general provisions (FAS-114) cannot be observed in call report data.

is not a capital model. Banks hold capital to cover not only credit losses but also to cover market losses, operational losses, and other losses. BankCaR is not designed to arrive at an appropriate level of capital for a given bank or for the banking system.

Probability model

This section describes the technical details of BankCaR's probability model. The bulk of the section describes the marginal distribution of the charge-off rate within a category. Once the marginal distributions are in place, they are readily connected to each other.

The underlying model is the structural credit portfolio model based on the insights of Robert Merton. The model produces the probability distribution of loss, relative to the return of par, resulting from default. It is commonly used to measure portfolio credit risk at banks. Recently, the model gained prominence as the basis for the Basel II minimum capital requirement.

Merton's model produces the distribution of the default rate. In addition, loss depends on the distribution of dollar exposure amounts and on the distribution of loss given default (LGD), which is the fraction of par that is lost on a defaulted instrument. BankCaR assumes that loan amounts are small enough that no single loan can have an appreciable effect on a charge-off rate. The role of LGD is developed later.

We begin with a particular loan made to a firm. The model assumes that when the loan matures there is a comparison between the firm's assets and liabilities. In the case of a surplus, the firm can refinance or find other ways to make payment. In the case of a shortfall, the firm defaults.

The model assumes that the value of the firm's liabilities is known and that the return on the firm's assets has a normal distribution. Standardizing the asset return has no effect on the idea that default occurs if and only if asset return is less than some fixed threshold. Symbolizing the standardized asset return as α_i:

(1) Firm i defaults if and only if $\alpha_i < \alpha_i^*$, where $\alpha_i \sim N[0,1]$

The probability that Firm i defaults is designated PD_i. Therefore, $PD_i = \Phi(\alpha_i^*)$, where $\Phi(\cdot)$ represents the standard normal cumulative distribution function (CDF). Using the inverse CDF, $\alpha_i^* = \Phi^{-1}(PD_i)$. In practice, PD refers to a fixed interval of time. In BankCaR, the period is one year, which is equal to the forecast horizon.

The asset returns of firms in the same category are assumed to be jointly normal. As such, they can be represented by a factor model. BankCaR assumes that there is a single systematic factor affecting the asset returns of all firms in the category.[3] Each firm's asset return is also affected by an independent idiosyncratic factor that is normally distributed. When the systematic and idiosyncratic factors are standardized, a single parameter, ρ_i, decomposes α_i into sources of risk:

(2) $$\alpha_i = \sqrt{\rho_i} Z + \sqrt{1-\rho_i} X_i; \quad Z, X_i \sim i.i.d.\ \mathrm{N}[0,1]$$

Z is the average standardized asset return of firms in the category and is referred to as the category factor. If Z takes a positive value, asset returns tend to be above average. Because X_i is independent of other information, there is no opportunity for Firm i to be affected by other firms that might share its industry, region, or other characteristic.

BankCaR assumes that all firms in the category are statistically identical. Thus, the parameters PD and ρ are uniform within the category. This implies that the correlation between the asset returns of two firms in a category is equal to ρ:

(3) $$Corr[\alpha_i, \alpha_h] = Corr[\sqrt{\rho} Z + \sqrt{1-\rho} X_i, \sqrt{\rho} Z + \sqrt{1-\rho} X_h] = \rho$$

Because of this, the parameter ρ in structural models is often referred to as "asset return correlation," and estimates of correlation based on asset returns are often employed in credit models. This puts great faith in the validity of the model's stylized assumptions. By contrast, BankCaR calibrates all parameters, including its correlations, to call report data. To reinforce this distinction, we refer to parameter ρ as "category correlation." Other things equal, a category with greater correlation would exhibit greater dispersion of the default rate.

The centerpiece of the model is the conditional probability that Firm i defaults. Supposing that the systematic factor Z takes the value z, the default of Firm i depends completely on the realization of its idiosyncratic factor X_i. This insight leads to an expression for the conditionally expected default rate that was first provided by Oldrich Vasicek:

(4) $$\begin{aligned} P[\alpha_i < \alpha_i^* \mid Z = z] &= P[\sqrt{\rho} z + \sqrt{1-\rho} X_i < \Phi^{-1}(PD)] \\ &= P\left[X_i < \frac{\Phi^{-1}(PD) - \sqrt{\rho}\, z}{\sqrt{1-\rho}}\right] = \Phi\left[\frac{\Phi^{-1}(PD) - \sqrt{\rho}\, z}{\sqrt{1-\rho}}\right] \end{aligned}$$

Expression (4) states the conditionally expected default rate for a firm. In an asymptotic portfolio, it is also the default rate for the category. If z is greater than zero the default rate is less than PD, while if z is somewhat less than zero the default rate is greater than PD.

[3] In the terminology introduced by Michael Gordy, within a category BankCaR is an asymptotic single risk factor model.

Allowing the subscript to distinguish categories and λ to represent the default rate, λ_j is an invertible function of the category factor:

$$\text{(5)} \qquad \lambda_j = \Phi\left[\frac{\Phi^{-1}(PD_j) - \sqrt{\rho_j}\, Z_j}{\sqrt{1-\rho_j}}\right]; \quad Z_j \sim N[0,1]$$

$$\text{(6)} \qquad Z_j = \frac{\Phi^{-1}(PD_j) - \sqrt{1-\rho_j}\,\Phi^{-1}(\lambda_j)}{\sqrt{\rho_j}} = z_j(\lambda_j)$$

The change-of-variable technique then provides the density of the category default rate:

$$\text{(7)} \qquad f(\lambda_j) = \left|\frac{d\, z_j(\lambda_j)}{d\,\lambda_j}\right| \phi(z_j(\lambda_j)) = \frac{\sqrt{1-\rho_j}}{\sqrt{\rho_j}} \frac{\phi\left(\dfrac{\Phi^{-1}(PD_j) - \sqrt{1-\rho_j}\,\Phi^{-1}(\lambda_j)}{\sqrt{\rho_j}}\right)}{\phi(\Phi^{-1}(\lambda_j))}$$

where $\phi(\cdot)$ represents the standard normal probability density function (PDF).

Figure 4.
Conditionally expected default rates

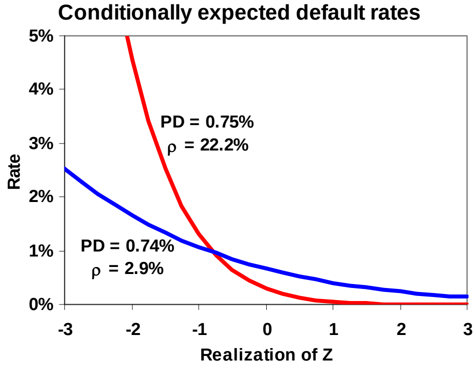
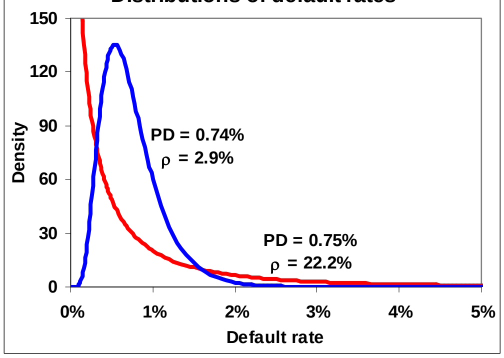


Figure 5.
Distributions of default rates



Figures 4 and 5 illustrate these functions for cases of high correlation (PD = 0.75%, ρ = 22.2%) and low correlation (PD = 0.74%, ρ = 2.9%).[4] Figure 4 shows that the two distributions are equally likely to produce default rates greater than about 1.0%. Figure 5 shows that the high-correlation distribution is much more likely to produce high rates of default. It should be apparent that this family of distributions can produce a large range of shapes, including highly skewed distributions.

To complete the specification of the probability model, BankCaR assumes that the standard normal category factors are ***jointly*** normal. Any pair of category

[4] These values reflect the estimates for construction and for lease financing, respectively.

factors is therefore connected by a factor correlation. Symbolizing the 12x12 matrix of factor correlations, $\Sigma = [[\rho_{ij}]]$, the joint PDF of the twelve default rates is given by change-of-variable:

$$(8) \qquad f(\lambda_1, \lambda_2, \dots, \lambda_{12}) = \phi_{12}(z_1(\lambda_1), z_2(\lambda_2), \dots , z_{12}(\lambda_{12}), \Sigma) \prod_{j=1}^{12} \frac{\sqrt{1-\rho_j}}{\sqrt{\rho_j}\ \phi(\Phi^{-1}(\lambda_j))}$$

where $\phi_{12}(\cdot, \Sigma)$ is the twelve-variable standard normal PDF with correlation matrix equal to Σ. The resulting connection between the marginal charge-off rate distributions is sometimes referred to as the Gauss copula.

Calibration

This section describes how BankCaR's parameters are calibrated to values that best reflect historical charge-off rates, according to the principle of maximum likelihood. The current values of the parameters appear in Table 2.

Four issues arise when calibration takes place. The first issue is that the theory presented in the previous section derives a distribution of the default rate. A default rate would be generally greater than a charge-off rate, because a bank does not lose 100% of the balance of every loan in every case. Call reports do not contain separate default rates and LGD rates. Lacking the data, BankCaR makes an identifying assumption to bridge the gap.

A naïve identifying assumption would be to assume that in any year the LGD rate equals the same fixed value. If this were 25%, a sequence of charge-off rates $\{CR_1, CR_2, \dots, CR_T\}$ would imply a sequence of default rates $\{CR_1/0.25, CR_2/0.25, \dots, CR_T/0.25\}$. The distribution of the default rate would be calibrated to these data, and the loss rate would equal one-quarter the default rate.

The naïve approach assumes that LGD does not vary with conditions, but the opposite has become widely accepted. The Basel II capital requirement, for example, recognizes that LGD can be greater in downturn conditions than in average conditions. If LGD rises at the same time as the default rate, there is more risk than if only the default rate rises in a downturn. For this reason, the naïve approach understates risk.

Instead of the naïve approach, BankCaR assumes that the risk of a loan depends on its expected loss and not on the decomposition of expected loss into PD and expected LGD. In particular, BankCaR assumes that risk would be unaffected if the unknown value of expected LGD were equal to 100%. This allows BankCaR to apply the theory of the default rate directly to the data on charge-off rates. Though for calibration purposes BankCaR assumes that LGD = 100%, there is no inference regarding the values of expected LGD and PD, which remain unknown. Implicitly, the annual average LGD varies in step with the default rate.

The second issue involves the maximization of the likelihood function. Many likelihood functions must be maximized in all parameters simultaneously; however, a two-stage approach finds the same results given the nature of the BankCaR likelihood function. In the first stage, BankCaR analyzes each of the twelve categories in turn. In a category $\boldsymbol{j}$, the observed charge-off rate in year $\boldsymbol{t}$ is $\boldsymbol{CR_{j,t}}$, and its unconditional expectation is $\boldsymbol{ECR_j}$. BankCaR assumes independence between years and maximizes the likelihood function derived from Expression (7):

$$(9)\quad \ln L(ECR_j, \rho_j; \{CR_{j,t}\}) = \prod_t \frac{\sqrt{1-\rho_j}}{\sqrt{\rho_j}} \frac{\phi\left(\dfrac{\Phi^{-1}(ECR_j) - \sqrt{1-\rho_j}\,\Phi^{-1}(CR_{j,t})}{\sqrt{\rho_j}}\right)}{\phi(\Phi^{-1}(CR_{j,t}))}$$

This provides maximum likelihood estimates of $\boldsymbol{ECR_j}$ and $\boldsymbol{\rho_j}$. These in turn imply estimates of $\boldsymbol{z_{j,t}}$ according to Expression (6). In the second stage, all twelve sequences $\{\hat{z}_{1,t}, \hat{z}_{2,t}, \ldots, \hat{z}_{12,t}\}$ are in hand, and correlations between them can be calculated as usual to produce maximum likelihood estimates of the factor correlations.

This two-stage approach produces exactly the same result as maximizing the joint likelihood function simultaneously in all ninety parameters. In the first stage, the parameters of categories are distinct. The parameters $\boldsymbol{ECR_1}$ and $\boldsymbol{\rho_1}$ do not appear in the marginal distributions of $\boldsymbol{CR_2}$, $\boldsymbol{CR_3}$, ..., or $\boldsymbol{CR_{12}}$. Therefore, those data tell nothing about the values of $\boldsymbol{ECR_1}$ and $\boldsymbol{\rho_1}$; the maximum likelihood estimates of $\boldsymbol{ECR_1}$ and $\boldsymbol{\rho_1}$ are the same whether they are obtained from the marginal likelihoods or from the joint likelihood. Since the same is true of every category, working category-by-category provides the same estimates of $\boldsymbol{ECR}$ and $\boldsymbol{\rho}$ as would be found using simultaneous estimation.

In the second state, when the category-by-category estimates of $\boldsymbol{ECR_j}$, and $\boldsymbol{\rho_j}$ are substituted into Expression (8), the only remaining unknowns appear in the correlation matrix Σ**:**

$$(10)\quad \ln L = \prod_t \phi_{12}(\hat{z}_1(CR_{1,t}), \hat{z}_2(CR_{2,t}), \ldots\ , \hat{z}_{12}(CR_{12,t}), \Sigma\ ) \prod_{j=1}^{12} \frac{\sqrt{1-\hat{\rho}_j}}{\sqrt{\hat{\rho}_j}\ \phi(\Phi^{-1}(CR_{j,t}))}$$

The maximum value of Expression (10) would therefore occur at the maximum likelihood estimates of the correlations between the set of inferred category factors $\{\hat{z}_1(CR_{1,t}), \hat{z}_2(CR_{2,t}), \ldots\ , \hat{z}_{12}(CR_{12,t})\}$. These are exactly the correlations that BankCaR produces in the second stage of the calibration.

Thus, the special nature of BankCaR's likelihood function—the lockstep relation between $\boldsymbol{CR}$'s and $\boldsymbol{Z}$'s, mediated by distinct $\boldsymbol{PD}$'s and $\boldsymbol{\rho}$'s—allows a two-step procedure to produce the same estimates as would be found using simultaneous estimation. In practice, the two-step procedure is superior because multidimensional routines can converge to points along constraints rather than to the interior global maximum found using the two-stage approach.

The third issue is that BankCaR departs from a common credit modeling practice that takes literally the role of asset correlation. In that practice, estimates of correlation based on asset return data are substituted into the credit risk model. The distributions that result only partly reflect credit loss data. In part, they also reflect the data used to calibrate asset return correlation. By contrast, the parameters of BankCaR reflect only the historical charge-off rates that have actually been experienced.

The fourth issue is that the charge-off series for the six real estate categories are not available prior to 1991. To handle this missing data, BankCaR employs a maximum likelihood technique provided by Donald Morrison. To improve the statistical fit in the 1984-1990 sub-period, BankCaR employs a 13x13 matrix of factor correlations that includes a factor for total real estate, for which charge-off rates are available for all years. Once the factor correlation estimates are in hand, no further use is made of the total real estate series.

Table 2. BankCaR parameter estimates and conditionally expected charge-off rates

	Category	ECR_j	ρ_j	CCR_j	*Factor correlations,* ρ_{ij}											
Not Real Estate	C&I	1.44%	4.2%	4.5%		-32%	46%	60%	83%	66%	59%	57%	29%	56%	-10%	84%
	Consumer	2.68%	2.3%	6.0%	-32%		-16%	-45%	-8%	-76%	-57%	-61%	-21%	-3%	1%	-63%
	Other Lending	1.23%	12.6%	7.7%	46%	-16%		79%	32%	6%	60%	51%	72%	66%	-6%	25%
	Depository Institutions	0.62%	26.8%	8.7%	60%	-45%	79%		39%	48%	65%	63%	58%	51%	-19%	45%
	Lease Financing	0.74%	2.9%	2.1%	83%	-8%	32%	39%		44%	34%	32%	9%	35%	-40%	60%
	Agriculture	1.00%	10.8%	5.9%	66%	-76%	6%	48%	44%		50%	56%	6%	11%	-20%	79%
Comm. R. E.	Construction	0.75%	22.2%	8.3%	59%	-57%	60%	65%	34%	50%		99%	85%	80%	13%	70%
	Non-farm non-residential	0.40%	10.6%	2.7%	57%	-61%	51%	63%	32%	56%	99%		81%	76%	10%	72%
	Multifamily	0.37%	15.5%	3.5%	29%	-21%	72%	58%	9%	6%	85%	81%		88%	17%	28%
	Farm	0.14%	2.3%	0.4%	56%	-3%	66%	51%	35%	11%	80%	76%	88%		23%	47%
Res. R.E.	1-4 Revolving	0.20%	0.7%	0.4%	-10%	1%	-6%	-19%	-40%	-20%	13%	10%	17%	23%		10%
	1-4 Other	0.15%	1.3%	0.4%	84%	-63%	25%	45%	60%	79%	70%	72%	28%	47%	10%	

Table 2 shows the BankCaR parameter estimates for 2007, using data through year-end 2006. Table 2 also shows the conditional charge-off rate (***CCR_j***) for each category. The ***CCR*** of a category is defined as the 99.5th percentile of the category charge-off rate. It comes about when the category factor is at its 0.5th percentile:

$$(11) \qquad CCR_j = \Phi\left[\frac{\Phi^{-1}(ECR_j) - \sqrt{\rho_j}\ \Phi^{-1}(0.005)}{\sqrt{1-\rho_j}}\right]$$

Among the factor correlations appearing in Table 2, most are positive, consistent with the commonplace observation that the economic cycle tends to affect most firms irrespective of category. The four commercial real estate factors are especially strongly related.

As we have seen, the identifying LGD assumption made by BankCaR infers greater risk than the naïve approach assuming a fixed value of LGD. The quantitative difference is not always great. Using the naïve approach with LGD assumed equal to 30%, the difference is less than 10% of CCR for ten of the twelve categories, and it is less than 20% of CCR in the remaining two categories. Thus, the effect of LGD variation in BankCaR is not extreme.

Confidence regions can be derived from the asymptotic likelihood ratio, which approaches the χ^2 distribution as the number of years rises. Taking separately each category, the following statistic has the $\chi_2{}^2$ distribution:

$$(12) \qquad -2\ Log\left[\frac{LnL\,(ECR_j = ecr_j,\ \rho_j = rho_j)}{Max\ LnL\,(ECR_j, \rho_j)}\right] \sim \chi_2^2$$

Figure 6. Parameter estimates, 95% confidence regions, and level curves of CCR at 1%, 5%, 10%, 20%, 30%, 40%

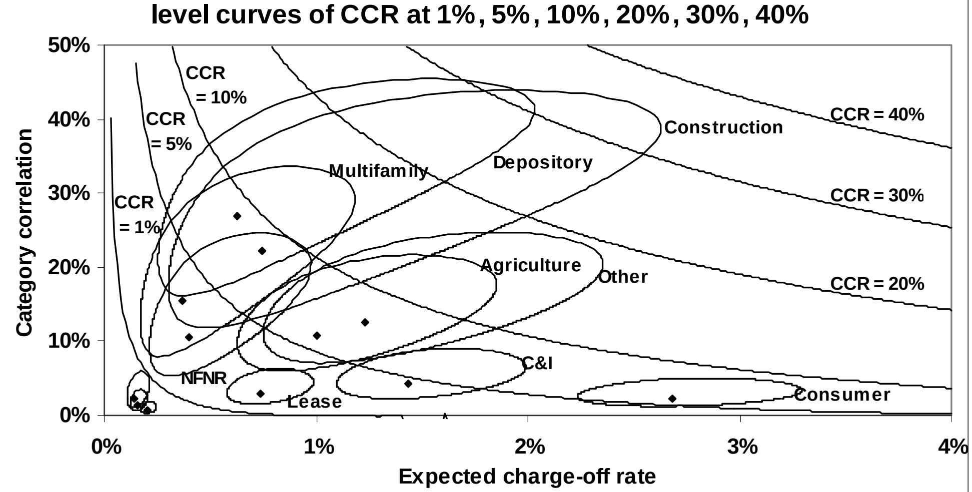


Figure 6 shows the maximum likelihood estimates of $\boldsymbol{ECR_j}$ and ρ_j and the 95% confidence regions derived from Expression (12). For the three categories appearing at the lower left (farm loans, one-to-four-family revolving residential mortgages, and one-to-four-family other residential mortgages) the confidence regions are small. For some other categories, the category-level data cannot reject a large range of combinations of $\boldsymbol{ECR}$ and ρ.

The parameter uncertainty implies uncertainty regarding conditional charge-off rates. Figure 6 illustrates this with level curves of CCR equal to 1%, 5%, 10%, 20%, 30%, and 40%. A separate indication is provided by bootstrapping. For three categories, we resample the charge-off data, re-estimate the parameters, and recalculate CCR 1,000 times. The bootstrapped 95% confidence intervals for CCR are 3.3% to 5.3% for C&I, 4.6% to 7.1% for consumer, and 1.4% to 18.5% for construction. The significant uncertainty in tail estimates is principally a consequence of the available data: we seek the worst charge-off that would come about once in 200 years, yet we have only sixteen years of construction charge-off data to work with. Especially given the pattern of the rates in certain categories, there is great uncertainty in projections of tail risk.

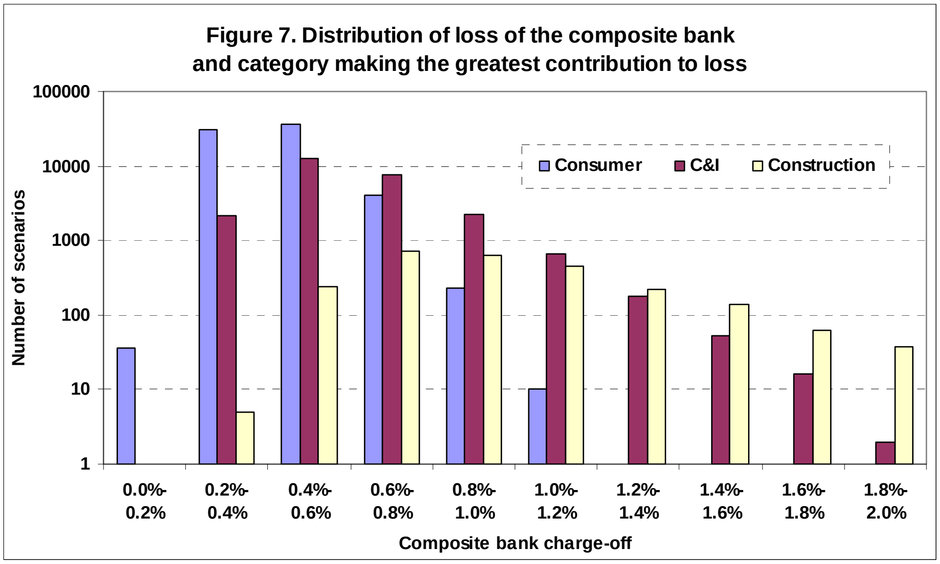
The US composite bank

This section shows some of the results that are available when BankCaR is applied to a bank. The bank in the spotlight is the US composite bank—a hypothetical bank representing the US banking sector as a whole. The loan amounts of the composite bank equal to the sums of loans at all commercial banks.

Table 3. Composite bank loss in an illustrative scenario

Lending Category	$Billions Outstanding	Illustrative scenario Charge-off rate	$B Charge-off
C&I	970	1.25%	12.1
Consumer	752	1.47%	11.0
Other Lending	176	1.33%	2.3
Depository Institutions	106	0.21%	0.2
Lease Financing	124	0.36%	0.4
Agriculture	54	0.71%	0.4
Construction	497	6.63%	33.0
Non-farm non-residential	808	2.11%	17.0
Multifamily	106	2.51%	2.7
Farm	52	0.28%	0.1
1-4 Revolving Residential	467	0.32%	1.5
1-4 Other Residential	1,430	0.21%	2.9
Non-Loan Assets	4,497	---	---
Totals	10,038		83.8
Composite bank loss in the illustrative charge-off scenario			0.83%

Table 3 displays the composite bank loan balances as of year-end 2006. Table 3 also shows an illustrative scenario that is an extension of an earlier example. The dollar charge-off in any category is the product of the loan balance and the charge-off rate. No charge-off results from assets other than loans. Combining charge-offs from all categories, the illustrative scenario produces a loss equal to 0.83% of assets. This would illustrate one of the 100,000 simulation runs that build up the distribution of a bank's loss distribution.



The distribution of loss for the composite bank is shown in Figure 7. Loss is highly concentrated in a narrow range: over 49% of scenarios produce loss in the range 0.40%-0.60% of assets. Because of this concentration, the vertical axis Figure 7 is shown on a logarithmic scale. Only one scenario in twenty produces loss greater than 0.80%. CaR is equal to the 500th greatest loss, 1.32%.

For each scenario, BankCaR determines the category that contributes the greatest amount of loss. Except in 0.04% of scenarios, the dominant category is consumer (71.8%), C&I (25.6%), or construction (2.6%). This breakdown parallels the breakdown of twenty-three years of historical data: consumer has been the greatest contributor in fifteen years, C&I in eight years, and construction in none of the years.

The breakdown is different, though, depending on the overall amount of loss. As shown in Figure 7, if overall loss is low or moderate, the consumer category or the C&I category is likely to produce the greatest loss. If overall loss is greater than about 1.20% the construction category becomes dominant. Thus, if the US banking system has a year of very high charge-offs overall, BankCaR projects that primary source of charge-offs will differ from what has been observed historically. An adverse year is likely to be dominated by construction charge-offs and more generally by the broader grouping of commercial real estate.

The average of the 1,377 scenarios producing greatest loss is the characteristic scenario for the composite bank. Not surprisingly, it projects a high rate of charge-off in the construction category.

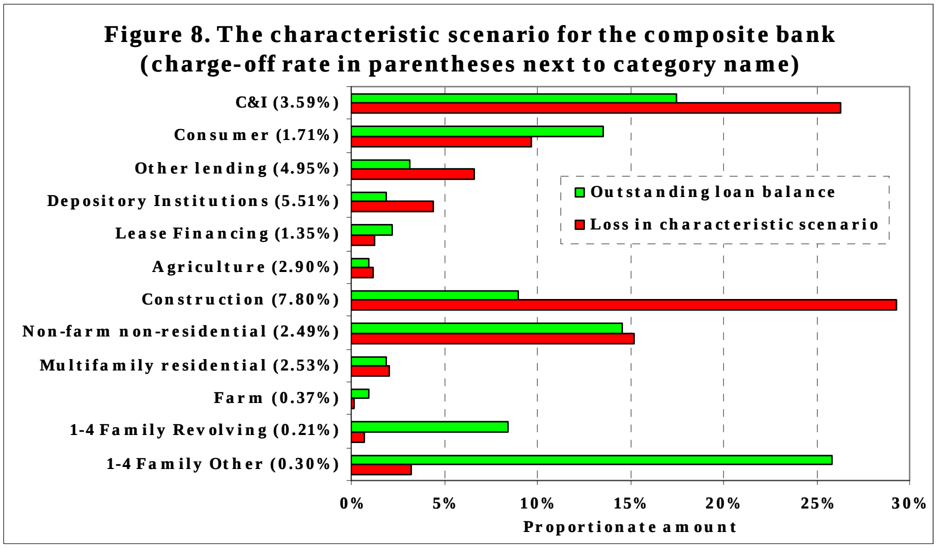


Figure 8 compares loan balances to loss amounts under the characteristic scenario of the composite bank. The characteristic scenario itself appears in parentheses next to the category names.

Though exposure to construction loans is only 9% of total loans, construction contributes more loss than any other category under the characteristic scenario. The composite bank is therefore said to have the construction risk type. C&I is a close second.

The composite bank enjoys a diversification benefit because the twelve category factors are correlated imperfectly. If instead we assume instead that all factors correlations equal 100%, every factor would attain its 99.5th percentile in the same scenario, and CaR would depend only on category ***CCRs***. For the composite bank, the loss would then equal 1.91%. The full-model value of CaR, 1.32%, therefore represents a diversification benefit equal to 30.8%.

Individual US Banks

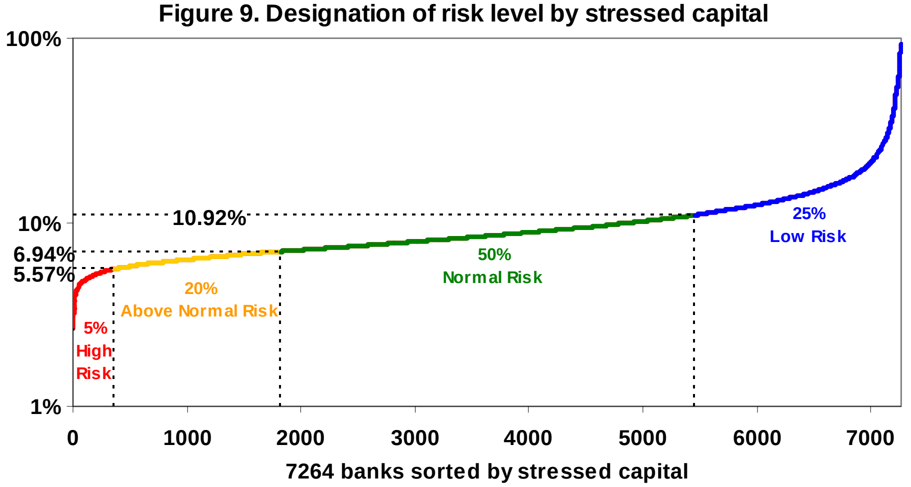
This section repeats the foregoing analysis for each of the 7264 commercial banks observed by BankCaR at year-end 2006. On average, the diversification benefit is 20.6%. This is naturally less than for the better-diversified composite bank. Among the individual banks, the diversification benefit ranges between 0% and 45%. Banks that specialize in a single category have the least diversification benefit. The greatest diversification benefit is enjoyed by banks having the consumer risk type because of the negative correlation between the consumer factor and other factors.

Table 4. US commercial banks by risk type		
RiskType	Number of banks	Average CaR
C&I	564	1.36%
Consumer	226	1.72%
Other	36	1.35%
Depository Institutions	17	1.18%
Lease Financing	5	1.45%
Agriculture	1,039	1.72%
Construction	4,921	2.01%
Non-farm non-residential	287	1.52%
Multifamily Residential	9	1.59%
Farm	1	1.39%
1-4 Family Revolving	1	0.12%
1-4 Family Other	158	1.33%
All Banks	7,264	1.87%

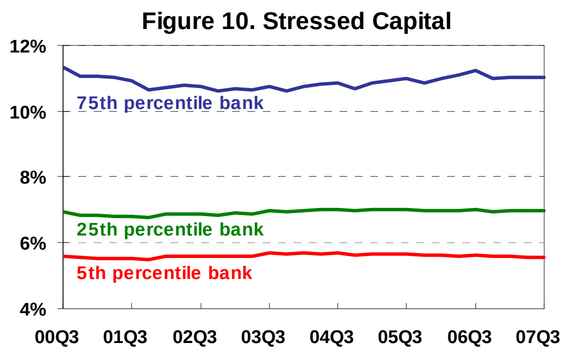
Table 4 shows the distribution of banks by risk type and the associated average value of CaR. Three of the risk types contain 90% of banks: C&I, agriculture, and construction. Some other risk types are poorly represented either because few banks have substantial exposure to the category (for example, lease financing), or because the CCR of the category is low (for example, farm lending).

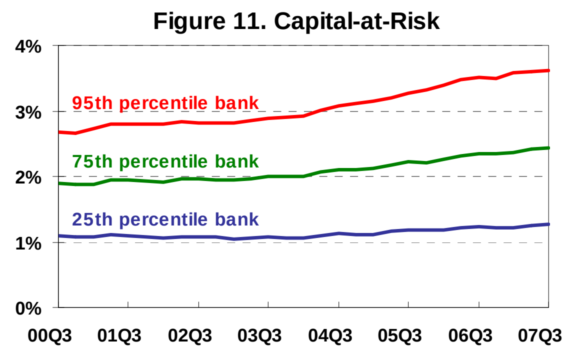
Construction, which has been shown to have outsized effects on the composite bank, also plays an important role at the individual bank level. Over 2/3 of banks have the construction risk type, and banks having the construction risk type have the greatest average value of CaR. An detailed analysis of the source of risk is available to regulators on a bank-by-bank basis as part of BankCaR's "dashboard" report.

Figure 9. Designation of risk level by stressed capital



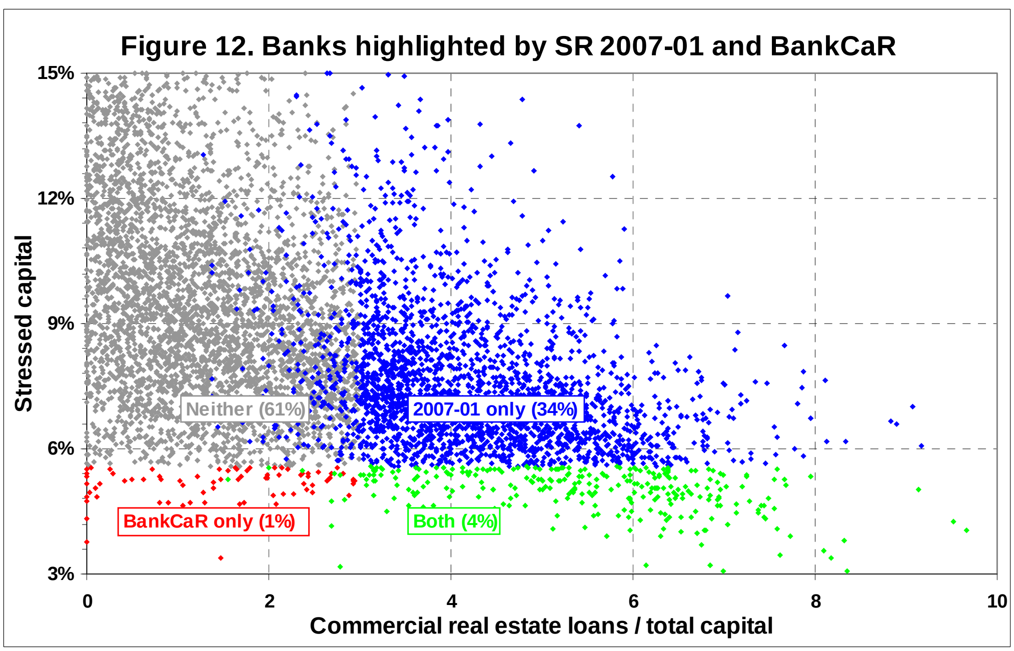
As a screening tool that compares credit risk between banks, BankCaR highlights banks that have low levels of stressed capital. Specifically, the banks having the lowest 5% of stressed capital are designated "High Risk", the next 20% are designated "Above Normal Risk", the center 50% are designated "Normal Risk", and the 25% with greatest Stressed Capital are designated "Low Risk". These designations are somewhat arbitrary in that bank risk is a continuous spectrum without definite break points, and yet a level of stressed capital at the risky end of the spectrum should not be ignored by bank managers or examiners. The current break points between designations are illustrated in Figure 9.





The break points are remarkably stable over time, as shown in Figure 10. This results from offsetting influences. The corresponding percentiles of CaR have risen over time, as shown in Figure 11. The general rise in CaR has been offset by a simultaneous rise in bank holdings of Tier 1 capital plus ALLL, producing stability in the distribution of stressed capital.

Figure 11 also shows that the inter-quartile range of CaR has increased over time. Credit risk at different banks differs more than it did previously. This trend toward diversity of risk increases the value of a risk-based tool such as BankCaR.



BankCaR's High Risk banks can be compared to the set of banks highlighted by an exposure-based screen. In Figure 12, each dot represents a US commercial bank. BankCaR would highlight banks lower in the diagram, that is, banks with relatively low values of stressed capital. To make this definite, the diagram isolates the 5% of banks that BankCaR designates as High Risk.

By contrast, an example of an exposure-based screen is Federal Reserve SR Letter 2007-01. This SR Letter highlights banks that have a high ratio of either CRE or construction lending to total capital.[5] This is a not unreasonable screening criterion, given that construction loans have statistically high risk as confirmed by the analysis of BankCaR. And yet, since it is based on exposure rather than on risk, it calls attention to banks having high levels of stressed capital. Those banks, recall, would have substantial levels of tier 1 plus ALLL even if they were to experience charge-off at the high level represented by CaR. An exposure-based criterion is also likely to miss some banks that have low levels of stressed capital, simply because those banks do not have the exposure that is being screened for.

A separate target of analysis is the set of BankCaR's scenarios. A particular scenario might cause a well capitalized bank to become adequately capitalized, undercapitalized, significantly undercapitalized, or critically undercapitalized.[6] Many BankCaR scenarios produce no change in the status of any bank, because they involve low and moderate rates of charge-off. Some scenarios produce a few status changes, and a few scenarios produce status changes at many banks.

Figure 13. Number of undercapitalized banks in the worst scenarios, with the category producing the greatest loss

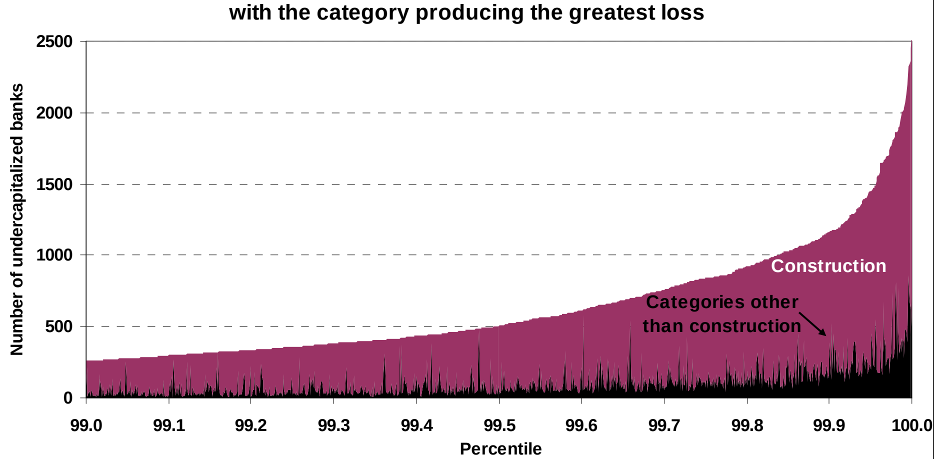


Figure 13 portrays scenarios that cause a well capitalized bank to become undercapitalized. Only 1,000 scenarios are shown, sorted by the number of banks affected. The worst scenario causes about 2,500 banks to become undercapitalized. Of these, about 1,800 would have greater charge-offs in construction than in any other category; the remaining 700 would have their greatest charge-offs in a category other than construction. All told, Figure 13

[5] SR 2007-01 also considers growth in the CRE portfolio and excludes lending on owner-occupied buildings. These conditions are not reflected in this comparison because of data limitations.
[6] These terms are defined in the FDIC Improvement Act of 1991, which places banks within categories defined in part by the ratio between capital and assets.

portrays 539,920 events in which a bank becomes undercapitalized. In 80% of those events, construction is the greatest source of loss.

The construction category has been encountered repeatedly in these analyses. When the composite bank has an unusually great loss, the category making the greatest contribution tends to be construction. Among individual banks, most have the construction risk type. Banks having the construction risk type tend to have a high level of CaR. In an adverse event, such as one of the scenarios portrayed in Figure 13, banks that become undercapitalized are likely to find that construction loans produce greater charge-offs than any other category. Despite all this, a bank's exposure to the construction category is not the whole story when it comes to forecasting its charge-off risk.

Conclusion

BankCaR forecasts the distribution of a commercial bank's charge-offs at the horizon of one year. This departs from traditional styles of forecasting that produce point estimates. One traditional style focuses on the asset quality of a bank's loans using tools such as regulatory classifications, the level and trend of loans past due, and other factors. Another style focuses on macroeconomic or financial variables such as output, interest rates, equity returns, or other data. Forecasts, for the bank or for the macro economy, are most accurate when the forecast horizon is short. To the extent that such forecasts are valid for a longer horizon, they could be incorporated into later versions of BankCaR.

Applied to the composite US commercial bank, BankCaR finds that the 99.5th percentile of the distribution of charge-offs is equal to the moderate value of 1.32%. If the US aggregate charge-off were this high, BankCaR predicts that construction lending would probably contribute more charge-offs than any other category. By contrast, the greatest contributor has historically been consumer or C&I lending. BankCaR predicts that a bad year would be dominated, for the first time, by construction.

Applied separately to each US bank, BankCaR efficiently identifies banks that have an adverse combination of credit risk and financial risk resources, defined as tier 1 capital plus ALLL. The difference, stressed capital, is a natural metric for comparing the inherent credit risk of banks. Over the last seven years, the distribution of stressed capital has been stable as a result of offsetting increases in risk and capital.

BankCaR uses publicly available regulatory reporting data, the most common credit portfolio model, and standard quantitative techniques. These generic qualities provide a standard of comparison between banks and a benchmark for more detailed models developed for specific banking institutions.

References

Federal Financial Institutions Examination Council (FFIEC), 2007, Consolidated Reports of Condition and Income, https://cdr.ffiec.gov/Public/WelcomeAdditionalInfo.aspx

Frye, Jon, 2000, "Collateral Damage," *Risk*, April, http://www.chicagofed.org/publications/publicpolicystudies/emergingissues/pdf/S&R_2000_15.pdf

Gordy, Michael, 2000, "A comparative anatomy of credit risk models," *Journal of Banking and Finance* 24, pages 119-149, http://www.federalreserve.gov/Pubs/feds/1998/199847/199847pap.pdf

Gordy, Michael, 2002, "A Risk-Factor Model Foundation for Ratings-Based Bank Capital Rules", *Journal of Financial Intermediation*, vol. 12 (July 2003), pp. 199-232, http://www.federalreserve.gov/pubs/feds/2002/200255/200255pap.pdf

McNeil, Alexander, Rudiger Frey and Paul Embrechts, 2005, *Quantitative Risk Management*, Princeton University Press

Merton, Robert, 1973, "On the Pricing of Corporate Debt: The Risk Structure of Interest Rates," *Journal of Finance*, Vol. 29, MIT (1974), pp. 449-470, http://dspace.mit.edu/bitstream/1721.1/1874/1/SWP-0684-14514372.pdf

Morrison, Donald F, 2005, *Multivariate Statistical Methods*, Fourth Edition. Brooks/Cole Publishing

Office of the Comptroller of the Currency, Federal Reserve System, Federal Deposit Insurance Corporation, and Office of Thrift Supervision, 2007, "Risk-Based Capital Standards: Advanced Capital Adequacy Framework," Basel II Preamble: http://www.occ.gov/ftp/release/2007-123a.pdf Basel II Final Rule: http://www.occ.gov/ftp/release/2007-123b.pdf

Vasicek, Oldrich, 2002, "Loan Portfolio Value," *Risk*, December, pp. 160-162, http://www.moodyskmv.com/conf04/pdf/papers/dist_loan_port_val.pdf

Working Paper Series

A series of research studies on regional economic issues relating to the Seventh Federal Reserve District, and on financial and economic topics.

Working Paper Series *(continued)*

Working Paper Series *(continued)*

Working Paper Series *(continued)*

Working Paper Series *(continued)*

Working Paper Series *(continued)*

The Spending and Debt Responses to Minimum Wage Increases *Daniel Aaronson, Sumit Agarwal, and Eric French*	**WP-07-23**
The Impact of Mexican Immigrants on U.S. Wage Structure *Maude Toussaint-Comeau*	**WP-07-24**
A Leverage-based Model of Speculative Bubbles *Gadi Barlevy*	**WP-08-01**
Displacement, Asymmetric Information and Heterogeneous Human Capital *Luojia Hu and Christopher Taber*	**WP-08-02**
BankCaR (Bank Capital-at-Risk): A credit risk model for US commercial bank charge-offs *Jon Frye and Eduard Pelz*	**WP-08-03**