The Expected Rate of Credit Losses on Banks' Loan Portfolios

Trevor S. Harris
Arthur J. Samberg Professor of Professional Practice
Columbia Business School
Email: tsh1@gsb.columbia.edu

Urooj Khan
Class of 1967 Associate Professor of Business
Columbia Business School
Email: uk2117@gsb.columbia.edu

Doron Nissim
Ernst & Young Professor of Accounting and Finance
Columbia Business School
Email: dn75@gsb.columbia.edu

The Accounting Review (Forthcoming 2018)

ABSTRACT

Estimating expected credit losses on banks' portfolios is difficult. The issue has become of increasing interest to academics and regulators with the FASB and IASB issuing new regulations for loan impairment. We develop a measure of the one-year-ahead expected rate of credit losses (*ExpectedRCL*) that combines various measures of credit risk disclosed by banks. It uses cross-sectional analyses to obtain coefficients for estimating each period's measure of expected credit losses. *ExpectedRCL* substantially outperforms net charge-offs in predicting one-year-ahead realized credit losses and reflects nearly all the credit loss-related information in the charge-offs. *ExpectedRCL* also contains incremental information about one-year-ahead realized credit losses relative to the allowance and provision for loan losses and the fair value of loans. It is a better predictor of the provision for loan losses than analyst provision forecasts and is incrementally useful beyond other credit risk metrics in predicting bank failure up to one year ahead.

Keywords: Banks; credit loss; loans; loan loss provisions; bank failure; analyst forecast; standard setting.

The authors gratefully acknowledge the helpful comments made by Mary Barth, Gauri Bhat, Leslie Hodder (editor), Scott Liao, Peter Raupach, Dushyant Vyas, Chris Williams, three anonymous reviewers, and seminar participants at the 2013 AAA Annual Meeting, the 24th Annual Conference on Financial Economics and Accounting, the 2014 joint workshop of Deutsche Bundesbank and the RTF of the Basel Committee on Banking Supervision, the Burton seminar at Columbia Business School, Emory University, London Business School, Singapore Management University and Washington University at St. Louis.

I. INTRODUCTION

This paper develops a measure of the expected rate of credit losses on a bank's loan portfolio using publicly available disclosures. For most banks, lending is the main source of value creation and risk, with economic profitability determined by the yield charged relative to cost of funds and credit risk realized. Accounting researchers have long studied the information contained in the various loan and related credit risk disclosures (e.g., Wahlen 1994; Barth et al. 1996; Nissim 2003; Khan and Ozel 2016) and in the wake of the 2007–2009 financial crisis, interest in the analysis of credit risk in banks has surged (e.g., Blankespoor et al. 2013; Cantrell et al. 2014). This interest goes beyond the academic literature. The Financial Accounting Standards Board (FASB) and the International Accounting Standards Board (IASB) have discussed how banks should report expected credit losses at the initiation of a loan. Their conclusions differ, but beginning in 2018, the IASB will require recognition of expected credit losses up to one year ahead at the initiation of a loan.[1]

To assess a bank's profitability and value, analysts, investors, and other users of financial statements who are not privy to private, highly disaggregated internal bank data need a good measure of expected credit losses that can be estimated from accounting disclosures and other public information.[2] For example, they may seek to independently estimate expected credit losses

[1] The IASB issued IFRS 9, *Financial Instruments*, on July 24, 2014, which requires recognition of expected credit losses. The new standard requires: "[A]t each reporting date, an entity would recognize a credit loss allowance or provision equal to 12-month expected credit losses (i.e., based on the probability of a default occurring in the next 12 months)" (Ernst and Young 2014: 6). IFRS 9 is effective for annual periods beginning on or after January 1, 2018. The FASB issued the current expected credit losses standard—ASU 2016-13—on June 16, 2016; this standard is effective for fiscal years beginning after December 15, 2019.

[2] In contrast, bank managers and bank examiners have private and disaggregated data to estimate expected credit losses. Auditors have access to similar data for the banks they audit but for comparative analysis, they, like other users, must rely on publicly available disclosures and other information.

to assess the quality of a bank's reported allowance for loan and lease losses (ALLL) and, by association, its provision for loan and lease losses (PLLL).

Banks publicly disclose information about loan yield, loan duration, and the composition of their loan portfolios, including the amount of nonperforming loans (NPLs). Each of these measures partly reveals credit quality. Banks also write off loans that are deemed to be uncollectible (charge-offs) and, when balance sheets are prepared, they report an ALLL that reflects a reserve for future write-offs of period-end loans. The ALLL is based on outstanding loan balances—and presumably on NPLs—but under current regulation, banks can only consider probable losses that can be estimated—rather than using an ex-ante notion of credit risk—to estimate the ALLL. The charge to income, the PLLL, increases the ALLL, which is reduced by net charge-offs (NCOs). Research has shown that each of these measures has shortcomings in measuring incurred credit losses.

We investigate whether existing credit-related measures and bank disclosures can be used together to better assess the next year's rate of realized credit losses. We formulate our measure—expected rate of credit losses, or *ExpectedRCL*—by estimating time-varying coefficients from cross-sectional regressions and then applying the coefficients to each bank's periodic measures of the relevant variables. We focus on the prediction of one-year-ahead credit losses because this is the period used in the estimation of the ALLL under certain regulatory guidance (such as that of the Federal Deposit Insurance Corporation (FDIC))[3] and in estimates of annual earnings and

[3] However, a bank can choose a different loss emergence period, depending on the composition of its loan portfolios, if it believes the losses will emerge over a different period. See https://www.fdic.gov/regulations/laws/rules/5000-4700.html.

profitability. A 12-month period is also the focus of immediate impairment recognition favored by the IASB under IFRS 9.[4]

We use accounting data from regulatory consolidated financial statements (FR Y-9C reports) for the period Q4:1996 through Q2:2015. The estimated coefficients on the variables included in our model of *ExpectedRCL* have the expected signs and are statistically significant. The most significant explanatory variables are NCOs, the level of NPLs, and a measure of unexpected change in NPLs. The coefficients generally keep the same sign throughout the sample period, but, as expected, almost all the magnitudes change significantly around the 2007–2009 financial crisis (henceforth, "the financial crisis"), consistent with a greater likelihood of credit losses in that period. We find that *ExpectedRCL* performs substantially better than NCOs in predicting one-year-ahead realized credit losses.

Throughout our sample period, a dollar of unexpected change in NPLs predicts substantially less than a dollar of credit losses. Also, the proportion of the unexpected change in NPLs that is equivalent to a credit loss that has yet to be charged off increased significantly during the financial crisis. This suggests that the increased credit losses during the crisis were not only due to the borrowers' deteriorating credit profiles but also to the greater loss implications of each dollar of NPLs and possibly to more aggressive charge-off policies.

Banks have disclosed the fair value of their loan portfolios since 1992 and, in concept, the fair value of a loan should reflect its *expected* credit and interest rate risks. Cantrell et al. (2014) compare the historical cost (net of the ALLL) and fair value of loans (FVLoans) to investigate which better reflects future credit losses. We extend this analysis by documenting that

---

[4] Further, moving beyond a year's horizon introduces significant measurement problems because of the turnover of loans and changing macro conditions, making controlling for other factors when using public disclosures much more difficult. Nonetheless, as an additional analysis, in Section VI we examine the predictive ability of *ExpectedRCL* for credit losses up to three years ahead.

ExpectedRCL contains incremental information relative to FVLoans in predicting one-year-ahead realized credit losses.

Next, we compare and contrast the forecasting ability of *ExpectedRCL* for next year's realized credit losses, relative to the ALLL and the PLLL. In standalone regressions, each dollar of *ExpectedRCL* translates, on average, into 96 cents of realized credit losses in the following 12 months, compared to only 41 cents for each dollar of the ALLL. *ExpectedRCL* also contains incremental information about one-year-ahead realized credit losses relative to the ALLL and the PLLL combined.[5]

To evaluate the generalizability of our results, we investigate the out-of-sample predictive ability of *ExpectedRCL* for one-year-ahead realized credit losses in the full sample and in subsamples based on bank size and loan portfolio composition. In all samples, *ExpectedRCL* has better predictive ability for one-year-ahead NCOs than ALLL, PLLL, NCOs, and FVLoans. The improvement in prediction offered by *ExpectedRCL* is economically meaningful as well. For example, relative to the ALLL, using *ExpectedRCL* to predict one-year-ahead NCOs reduces the absolute prediction error for the average bank by 24%. This suggests that *ExpectedRCL* is a better predictor of one-year-ahead realized credit losses than other publicly disclosed credit-risk-related metrics.

We conduct additional analyses to establish the usefulness of *ExpectedRCL* and test the robustness of our findings. First, we document that *ExpectedRCL* has better predictive ability for the PLLL relative to analysts' PLLL forecasts. Further, on average, banks have larger earnings surprises the greater the differences between *ExpectedRCL* and analysts' PLLL forecasts or

[5] Not surprisingly, *ExpectedRCL* does not subsume all the information in the ALLL and the PLLL, as these are based on more detailed inputs and can reflect private information available to managers.

between *ExpectedRCL* and forecasts of one-year-ahead realized credit losses based on ALLL, PLLL, or NCOs. Second, *ExpectedRCL* is incrementally useful beyond other credit risk metrics in predicting bank failure up to one year ahead. Finally, we examine whether *ExpectedRCL* can be used to predict credit losses beyond one year. Using the subsequent three-year net charge-off rate as a measure of long-horizon credit losses, we find that *ExpectedRCL* continues to display significant incremental information in the prediction of long-term credit losses relative to ALLL, PLLL, NCOs, and FVLoans.

By providing a more predictive measure of expected credit losses, our study contributes to research in accounting, banking, and finance. Our findings are relevant for the literature that explores whether accounting disclosures provide useful information about future credit losses. Past studies (e.g., Cantrell et al. 2014) have used net charge-offs or nonperforming loans as measures of credit risk. Our metric, *ExpectedRCL*, better estimates one-year-ahead realized losses on banks' loan portfolios. It can also be used as an additional explanatory variable in models predicting bank failure or earnings surprises.

Our measure has practical applications. First, since we find that *ExpectedRCL* is a better predictor of expected credit losses than the ALLL or the PLLL, analysts and investors might consider using *ExpectedRCL*, instead of—or in addition to—ALLL or PLLL to better assess banks' one-year-ahead realized credit losses. Second, investors and regulators might use it to identify banks in which the difference between reported PLLL and *ExpectedRCL* is among the largest in the cross-section or is deviating from past patterns.[6] Third, the recently issued IFRS 9, *Financial Instruments*, requires entities to recognize 12-month expected credit losses on their loan

[6] This is a commonly used approach by investors who conduct comparative quantitative analysis as part of their investment decisions.

portfolios at the initiation of loans. *ExpectedRCL* can serve as a benchmark to compare with those disclosures.

In summary, *ExpectedRCL* is a particular combination of publicly available credit risk disclosures of banks that outperforms other publicly disclosed credit risk metrics in predicting one-year-ahead credit losses. We do not, however, claim that it is the *optimal* measure. Due to the richness of detailed bank disclosures, other summary statistics can be constructed using linear and nonlinear combinations that may predict credit losses even better.

The rest of the study proceeds as follows. Section II discusses credit-risk-related measures disclosed by banks. Section III develops the methodology for estimating *ExpectedRCL*. Section IV discusses the sample selection procedures and sample data. Section V presents empirical findings. Section VI provides additional analyses and robustness tests. Section VII concludes the study.

II. PUBLICLY DISCLOSED METRICS RELEVANT TO A STRUCTURAL MODEL OF CREDIT RISK

Interest income is recognized over time and is derived from a yield that includes at least four components: the time-value of money, expected credit losses, risk premia, and economic profit. Because measuring expected losses is particularly complex, the timing of loss recognition is controversial and is frequently debated by regulators and practitioners. Under current US GAAP, credit losses for loans measured at amortized cost are based primarily on SFAS 5, *Accounting for Contingencies* (Accounting Standards Codification (ASC) subtopic 450-20), for unimpaired loans and on SFAS 114, *Accounting by Creditors for Impairment of a Loan* (ASC subtopic 310-10), for impaired loans.[7] SFAS 5's recognition criteria require that credit losses be probable and that they can be reliably estimated; such losses are usually referred to as incurred losses. Credit losses on portfolios of individually small and homogeneous unimpaired loans (e.g., residential real estate loans, credit card receivables, and other consumer loans) are usually estimated using statistical models based on historical data and annualized past experience. On the other hand, credit losses on individually large and heterogeneous unimpaired loans (e.g., commercial and industrial loans) are typically evaluated on a loan-by-loan basis. For both types of loans, the ALLL reflects the bank's estimate of probable losses based on events that have occurred up to that time rather than all expected future losses. In contrast, for impaired loans, the related ALLL does include some expected future losses. SFAS 114 considers a loan impaired when it is probable that the full contractual payments will not be received. For specific impaired loans, SFAS 114 generally requires that the ALLL be increased so as to reduce the net book value of the loans to the present value of expected cash receipts calculated using the effective interest rate. Still, in most cases, the portion of the total ALLL related to expected future credit losses (as opposed to incurred losses) is relatively small.

Importantly, the ALLL varies with the composition of the loan portfolio itself as well as with the relative conservativeness of any charge-off policy adopted by the management. Any charge-offs impact the loan balances, too. The PLLL is measured as the total of net charge-offs and the change in the ALLL due to operating activities. It thus includes, in part, (a) credit losses attributable to loans originating during the year and (b) any measurement errors in either the beginning or ending ALLL.

[7] Under international accounting standards, the accounting is based on IAS 39, *Financial Instruments: Recognition and Measurement*, subject to the changes in IFRS 9, *Financial Instruments*, issued in July 2014 and effective in 2018.

Although users of banks' financial information often use the ALLL and the PLLL as indicators of credit risk or expected credit losses,[8] these metrics have limitations. First, both are discretionary. Research shows that banks have used discretion in estimating the ALLL and the PLLL to signal private information as well as to manage book value, earnings, regulatory capital, and taxes.[9] Second, even in the absence of intentional bias, the ALLL and the PLLL are subjective estimates of future events. Third, the ALLL and the PLLL, under current accounting rules, do not reflect all expected losses that might be anticipated at the inception of the loan and priced into the yield. As stated in the FDIC's Interagency Policy Statement (Federal Deposit Insurance Corporation 2006: p. 3): "Under GAAP, the purpose of the allowance for loan and lease losses is not to absorb all of the risk in the loan portfolio, but to cover probable credit losses that have already been incurred."

The ALLL and the PLLL are estimated by managers based in part on a series of *primary* indicators, many of which are available in public disclosures. We focus on these primary indicators in constructing an alternative summary measure of the expected rate of credit losses.

Loan Balances and Loan Composition

Characteristics of the borrower and of the collateral, including the location of both, and the duration of the loan (especially as this relates to business cycles) affect both the probability of default and the loss-given-default. Some of these factors can be captured by examining the different types of loans making up the aggregate loan portfolio. Generally, banks' loan portfolios consist of real estate (the largest group), commercial and industrial (C&I), consumer, and other

[8] Typical ratios reported in analysts' reports include PLLL/Average Loans, ALLL/Loans, and NCOs/Loans (Ryan 2007).

[9] For example, see Beaver et al. 1989; Moyer 1990; Elliott et al. 1991; Wahlen 1994; Beatty et al. 1995; Collins et al. 1995; Beaver and Engel 1996; Ahmed et al. 1999; Liu and Ryan 2006; Bushman and Williams 2015.

loans. The "other" category includes lease financing receivables and loans to depository institutions, farmers, nondepository financial institutions, foreign governments, and official institutions. Given the differences in loss emergence for different loan types, our model of *ExpectedRCL* includes the proportions of the three largest loan categories: real estate, consumer, and other loans (which includes C&I loans).

Loan Yield

Because finance theory suggests that expected losses are priced into the yield, our model of *ExpectedRCL* includes loan yield. However, loan yield is not a perfect proxy for credit risk, as it also reflects interest rate risk and other risks and provisions (such as prepayment risk and call options). Thus, loan yield—measured as interest income divided by the average balance of loans—is a noisy measure of expected losses at the inception of the loan.

Loan Duration

In many cases, the longer the loan horizon, the more uncertainty there is about the underlying business (e.g., potential for default). On the other hand, the bank's willingness to extend a long-term loan depends on the perceived stability of the borrower. Either way, loan duration may provide relevant information about expected credit losses.

Nonperforming Loans

Loans that are not paying interest or principal due to a borrower's credit problems are classified as nonperforming loans and are an obvious factor to use in a model of expected credit losses. NPLs include nonaccrual loans, restructured (troubled) loans, and some past-due loans.

NPLs are considered relatively nondiscretionary (Beaver et al. 1989; Griffin and Wallach 1991) and prior studies have therefore used them as instruments to partition other measures of credit quality into discretionary and nondiscretionary components (Wahlen 1994; Collins et al.

1995; Beaver and Engel 1996). Beaver et al. (1989) indicate that, although nonaccrual and restructured loans are relatively nondiscretionary, their measurement does involve judgment that varies across banks. The impact of that discretion can be mitigated by redefining NPLs to include accruing loans that are at least 90 days delinquent. The probability of default and the expected loss given default vary substantially across loan categories of NPLs; for example, because of collateral guarantees by the United States government or its agencies (Araten et al. 2004).

Net Charge-offs

Net charge-offs (NCOs) are measures of realized credit loss in a given period and indirectly impact the balance sheet and income statement through the ALLL and the PLLL, respectively. NCOs have been used as a measure of credit risk in prior research (e.g., Cantrell et al. 2014) and in recent analyses of top-down stress tests (e.g., Hirtle et al. 2015), in part because they are considered relatively nondiscretionary (Moyer 1990; Wahlen 1994; Collins et al. 1995; Beaver and Engel 1996).[10] However, some discretion remains and prior studies demonstrate discretionary charge-offs by banks for earnings management (e.g., Liu and Ryan 2006). The discretion available to managers can also be used to signal and convey their detailed and disaggregated private information about the condition of loans.

III. METHODOLOGY

Given the above discussion, we specify the following model for the expected rate of credit losses for firm i in period t, based on information available at time t-1 ($ExpectedRCL_{i,t-1}$):

[10] The likely reason is that regulatory policies require banks to charge off particular loans when they have been delinquent for a certain number of days. For example, closed-end retail loans that become past due 120 cumulative days and open-end retail loans that become past due 180 cumulative days from the contractual date should be charged off.

$$ExpectedRCL_{i,t-1} = \propto_{0,t} + \propto_{1,t} RealizedRCL_{i,t-1} + \propto_{2,t} \frac{NPL_{i,t-1}}{Loans_{i,t-1}} + \propto_{3,t} LoansYield_{i,t-1} + \propto_{4,t} FloatLoanRatio_{i,t-1} + \propto_{5,t} \frac{RELoans_{i,t-1}}{Loans_{i,t-1}} + \propto_{6,t} \frac{ConsLoans_{i,t-1}}{Loans_{i,t-1}} + \varepsilon_{i,t-1} \quad (1)$$

where $RealizedRCL_{i,t-1}$ is the realized rate of credit losses of firm i in period t−1, measured relative to the average balance of loans during that period. We define this variable more precisely as we develop the model. $Loans_{i,t-1}$ is the total of loans held for investment of firm i at time t-1. $NPL_{i,t-1}$ is nonperforming loans of firm i at time t-1. As discussed previously, *NPL* is defined as the total of non-accruing loans, restructured loans, and accruing loans that are at least 90 days delinquent. $LoansYield_{i,t-1}$ is firm i's ratio of tax-equivalent interest income on loans to the average balance of loans over period t−1. $FloatLoanRatio_{i,t-1}$ is an estimate of the proportion of loans of firm i at time t−1 that reprice or mature within one year; it serves as a proxy for loan duration.[11] $RELoans_{i,t-1}$ and $ConsLoans_{i,t-1}$ are loans of firm i at time t−1 classified as real estate loans and consumer loans, respectively. The intercept ($\propto_{0,t}$) and coefficients ($\propto_{5,t}$ and $\propto_{6,t}$) of the two loan composition variables ($RELoans_{i,t-1}$ and $ConsLoans_{i,t-1}$) capture the average effects of the three primary loan categories. $\varepsilon_{i,t-1}$ represents the net effect of all other relevant information at time t-1 for the prediction of firm i's rate of credit losses in period t that is omitted from Equation (1). The definitions and details of the estimation of the variables included in our analysis are provided in Appendix A. The subscript of the model's coefficients is time t because, as explained below, these

[11] Specifically, we estimate the proportion of floating-rate loans using the ratio of floating-rate loans and securities to the total of loans and securities. We estimate floating-rate loans and securities by subtracting the total of (a) interest-bearing balances due from depository institutions and (b) federal funds sold and securities purchased under agreements to resell from earning assets that can be repriced within one year or that mature within one year.

coefficients are estimated in a regression in which the dependent variable is based on information at time t.

Since our main objective is to build an estimate of expected credit losses using a structural model that can, in part, be used to benchmark and assess the quality of the ALLL and the PLLL, we exclude those two from Equation (1). Instead, we use them to validate the performance of *ExpectedRCL* in predicting credit losses and use other primary credit-related measures in our structural model.

Equation (1) cannot be directly estimated because *ExpectedRCL* is unobservable. However, with unbiased expectations, the difference between the realized and expected rate of credit losses in period t should be unpredictable white noise:

$$RealizedRCL_{i,t} = ExpectedRCL_{i,t-1} + \varepsilon_{i,t}^{RCL} \tag{2}$$

Thus, Equation (1) can be reexpressed by substituting Equation (2) into Equation (1):

$$\begin{aligned} RealizedRCL_{i,t} \\ &= \propto_{0,t} + \propto_{1,t} RealizedRCL_{i,t-1} + \propto_{2,t} \frac{NPL_{i,t-1}}{Loans_{i,t-1}} \\ &+ \propto_{3,t} LoansYield_{i,t-1} + \propto_{4,t} FloatLoanRatio_{i,t-1} \\ &+ \propto_{5,t} \frac{RELoans_{i,t-1}}{Loans_{i,t-1}} + \propto_{6,t} \frac{ConsLoans_{i,t-1}}{Loans_{i,t-1}} + \varepsilon_{i,t}^{+} \end{aligned} \tag{3}$$

where $\varepsilon_{i,t}^{+} = \varepsilon_{i,t-1} + \varepsilon_{i,t}^{RCL}$. $RealizedRCL_{i,t}$ can only be observed after period t and is yet to be precisely defined.

To measure *RealizedRCL*, we start with NCOs during the period. Ideally, we seek a nondiscretionary measure of economically required charge-offs. NCOs can be relatively untimely, especially for large heterogeneous loans (e.g., C&I loans) and when economic conditions are changing (Ryan 2007). Banks may also delay charging off loans to avoid a decline in the ALLL that leads to an increase in the PLLL (e.g., Vyas 2011; Calomiris and Nissim 2014). Therefore, to derive a less-discretionary estimate of realized credit losses, we use the relationship between loans, NPLs, and NCOs to estimate and undo the discretionary component of NCOs. A discretionary NCO policy affects the level of NPLs, as discretionary acceleration of NCOs leads to lower NPLs, whereas slowing the rate of NCOs can leave loans as NPLs. Specifically, we estimate the unexpected change in NPLs during a period and use the negative of a fraction of the unexpected change in NPLs to estimate the discretionary charge-offs (we estimate discretionary charge-offs using only a portion of the unexpected change in NPLs because not all NPLs become NCOs). That fraction also captures the credit-loss equivalent of unexpected changes in NPLs that arises from changes in the credit quality of the loan portfolio or from changes in macroeconomic conditions.

To estimate the unexpected change in NPLs, we recognize that, when macroeconomic conditions and the credit quality of loans are relatively stable, changes in NPLs should stem from changes in the size of the loan portfolio. Thus, any increase in NPLs that cannot be attributed to a change in the size of the loan portfolio suggests that either (a) the macroeconomic conditions and/or the credit quality of the loan portfolio has changed during the period or (b) the bank has misstated its NCOs. Either way, to derive a less discretionary measure of realized credit loss, we need to adjust current NCOs for the portion of the unexpected change in NPLs unrelated to the change in the size of the loan portfolio. We start by estimating the unexpected change in NPLs during period *t* (ΔNPL_t^{unexp}) as:

$$\Delta NPL_{i,t}^{unexp} = NPL_{i,t} - Loans_{i,t} \times \frac{NPL_{i,t-1}}{Loans_{i,t-1}} \tag{4}$$

and specify the realized rate of credit losses as:

$$RealizedRCL_{i,t} = \frac{NCO_{i,t} + \gamma_t \Delta NPL_{i,t}^{unexp}}{AveLoans_{i,t}} = \frac{NCO_{i,t}}{AveLoans_{i,t}} + \gamma_t \frac{\Delta NPL_{i,t}^{unexp}}{AveLoans_{i,t}} \tag{5}$$

where $NCO_{i,t}$ is net charge-offs for firm i in period t. γ_t is a cross-sectional constant that is estimated each quarter and varies over time. $\gamma_t \Delta NPL_{i,t}^{unexp}$ represents the amount of unexpected change in NPLs that is equivalent to a credit loss that the bank has yet to charge off, either because the loss recognition criteria have not been met or because management has used its discretion to understate charge-offs. Thus, adding this amount to $NCO_{i,t}$ should result in a more complete measure of credit losses for bank i in period t. $AveLoans_{i,t}$ is the average balance of loans held by firm i during period t.

Using Equation (5), Equation (3) can be reexpressed as follows:

$$\frac{NCO_{i,t}}{AveLoans_{i,t}} = \propto_{0,t} + \propto_{1,t} \frac{NCO_{i,t-1}}{AveLoans_{i,t-1}} + \propto_{1,t} \gamma_t \frac{\Delta NPL_{i,t-1}^{unexp}}{AveLoans_{i,t-1}} - \gamma_t \frac{\Delta NPL_{i,t}^{unexp}}{AveLoans_{i,t}}$$

$$+ \propto_{2,t} \frac{NPL_{i,t-1}}{Loans_{i,t-1}} + \propto_{3,t} LoansYield_{i,t-1} + \propto_{4,t} FloatLoanRatio_{i,t-1} \qquad (6)$$

$$+ \propto_{5,t} \frac{RELoans_{i,t-1}}{Loans_{i,t-1}} + \propto_{6,t} \frac{ConsLoans_{i,t-1}}{Loans_{i,t-1}} + \varepsilon_{i,t}^{+}$$

Equation (6) can be estimated because at the end of period t, all the variables are observable, including $\frac{\Delta NPL_{i,t}^{unexp}}{AveLoans_{i,t}}$. However, OLS estimation would result in biased and inconsistent estimates because $\frac{\Delta NPL_{i,t}^{unexp}}{AveLoans_{i,t}}$ is likely to be strongly positively correlated with $\varepsilon_{i,t}^{+}$. This is because unexpected shocks to credit quality are likely to affect both NPLs and realized credit losses. Fortunately, consistent estimates of the parameters can still be derived by redefining the intercept and disturbance of Equation (6) as follows:

$$\propto_{0,t}{}^{*} = \propto_{0,t} - \gamma_t \frac{\overline{\Delta NPL_t^{unexp}}}{AveLoans_t} \qquad (7)$$

$$\varepsilon_{i,t}^{*} = \varepsilon_{i,t}^{+} - \gamma_t \left(\frac{\Delta NPL_{i,t}^{unexp}}{AveLoans_{i,t}} - \overline{\frac{\Delta NPL_{t}^{unexp}}{AveLoans_{t}}} \right) \tag{8}$$

where $\overline{\frac{\Delta NPL_{t}^{unexp}}{AveLoans_{t}}}$ is the cross-sectional average of $\frac{\Delta NPL_{i,t}^{unexp}}{AveLoans_{i,t}}$. We therefore estimate the following model for each quarter t:

$$\begin{aligned} \frac{NCO_{i,t}}{AveLoans_{i,t}} = \propto_{0,t}{}^{*} + \propto_{1,t} \frac{NCO_{i,t-1}}{AveLoans_{i,t-1}} + \propto_{1,t} \gamma_t \frac{\Delta NPL_{i,t-1}^{unexp}}{AveLoans_{i,t-1}} \\ + \propto_{2,t} \frac{NPL_{i,t-1}}{Loans_{i,t-1}} + \propto_{3,t} LoansYield_{i,t-1} \\ + \propto_{4,t} FloatLoanRatio_{i,t-1} + \propto_{5,t} \frac{RELoans_{i,t-1}}{Loans_{i,t-1}} \\ + \propto_{6,t} \frac{ConsLoans_{i,t-1}}{Loans_{i,t-1}} + \varepsilon_{i,t}^{*} \end{aligned} \tag{9}$$

Equation (9) satisfies the OLS assumptions because, by definition, unexpected shocks to realized credit losses of firm i at time t (i.e., $\varepsilon_{i,t}^{*}$) are uncorrelated with time t−1 information, as measured by the explanatory variables. The adjustment to the intercept is required because, in any given period, the average credit loss shock across all banks is not likely to be zero. However, because this adjustment is assumed to be constant in the cross-section, it does not affect the cross-sectional differences in the estimated rate of credit losses across banks (on which we focus).[12]

For each quarter during the sample period, we estimate cross-sectional regressions of Equation (9) using the trailing four quarters of data. We then use Equation (7) to estimate the

[12] It is counterintuitive that one can eliminate bias by excluding a variable and, indeed, in most cases, omitting a variable would introduce or increase the bias of the remaining coefficients. However, our case is unique in that the regression coefficients are related to each other (through γ and $\propto_1$). To see a simpler example of the same effect, assume that both X1 and X2 affect Y but X2 is correlated with the disturbance. Assume further that X1 is uncorrelated with either X2 or the disturbance and that its effect on Y is the same as that of X2. The full model, i.e., Y = a0 + a1 X1 + a1 X2 + e, which incorporates the restriction that the coefficients on X1 and X2 should equal each other, would result in a biased estimate of a1 because X2 is correlated with the disturbance. The reduced model, i.e., Y = a0 + a1 X1 + e, would result in an unbiased estimate of a1 because X1 is uncorrelated with the disturbance.

intercept, $\propto_{0,t}$, and calculate the expected rate of credit losses for the next year ($\widehat{ExpectedRCL_{i,t}}$) using the estimated parameters and the current (time t) values of the explanatory variables for each firm i:

$$\begin{aligned}
&\widehat{ExpectedRCL_{i,t}} \\
&\quad = \hat{\alpha}_{0,t} + \hat{\alpha}_{1,t}\frac{NCO_{i,t}}{AveLoans_{i,t}} + \hat{\alpha}_{1,t}\hat{\gamma}_t\frac{\Delta NPL_{i,t}^{unexp}}{AveLoans_{i,t}} + \hat{\alpha}_{2,t}\frac{NPL_{i,t}}{Loans_{i,t}} \\
&\quad + \hat{\alpha}_{3,t}LoansYield_{i,t} + \hat{\alpha}_{4,t}FloatLoanRatio_{i,t} + \hat{\alpha}_{5,t}\frac{RELoans_{i,t}}{Loans_{i,t}} \qquad (10) \\
&\quad + \hat{\alpha}_{6,t}\frac{ConsLoans_{i,t}}{Loans_{i,t}}
\end{aligned}$$

$\widehat{ExpectedRCL_{i,t}}$ is our estimate at time t of next year's (t+1) expected rate of credit losses on bank i's portfolio of held-for-investment loans.[13] In concept, $\widehat{ExpectedRCL}$ could incorporate more disaggregated classifications of some of the measures publicly disclosed by banks. However, many of the important inputs in our structural model—such as interest income (used to calculate loan yield) and loan maturity data—are only available at the aggregated-loan-portfolio level. Moreover, several of the variables used to validate the predictive ability and information content of $ExpectedRCL$—including ALLL, PLLL, and fair value of loans—are only available at the aggregated-loan-portfolio level.[14]

[13] Under current accounting rules, realized losses in a given period can reflect three types of losses – (1) losses incurred as of the previous measurement date and settled in the current period; (2) losses occurring and being realized in the current period that did not exist as of the measurement date; and (3) losses that were expected as of the measurement date but did not meet the probability threshold for recognition. Given that realized losses include losses in category (3), we refer to our measure as *ExpectedRCL*, a metric that provides an estimate of one-year-ahead realized losses.

[14] Beginning in 2008, interest income data becomes available for the following subcategories of loans: loans secured by one to four family residential properties, other loans secured by real estate, and all other loans. Also, the ALLL is reported at the aggregated-loan-portfolio level in the FR Y-9C reports until 2013.

IV. SAMPLE AND DATA

We focus on bank holding companies (BHCs) and extract accounting data from regulatory consolidated financial statements (FR Y-9C reports) for the period Q4:1996–Q2:2015. BHCs with total consolidated assets above $150 million or those that satisfy certain other conditions (e.g., those with public debt) were required to file the FR Y-9C report quarterly through the fourth quarter of 2005. The asset-size threshold was increased to $500 million in March 2006 and to $1 billion in March 2015. To make the sample comparable over time, we delete observations with total assets less than $1 billion at March 2015 prices. Our results are not sensitive to using $500 million in March 2006 prices as the cutoff instead.

We start the sample period in 1996 because information required for measuring certain FR Y-9C variables is unavailable before then. We measure all income statement quantities using the trailing four quarters of data to eliminate the effects of seasonality and to smooth out short-term shocks.[15] Thus the sample includes 75 quarters of data (Q4:1996 through Q2:2015). To mitigate the impact of outliers, we trim extreme values of each variable.[16] Summary statistics from the distributions of the trimmed variables are provided in Panel A of Table 1. For our sample, the mean (median) ALLL is 1.58% (1.41%) of gross loans held for investment. The ratio of the PLLL to average gross loans has a mean (median) of 0.62% (0.35%). On average, 1.77% of gross loans are classified as NPLs and 0.08% are estimated to be unexpected ΔNPLs. The mean (median)

---

[15] Seasonality affects quarterly data for accounting as well as economic reasons. For example, Liu et al. (1997) find that loan provisions are often delayed to the fourth fiscal quarter, when the audit occurs.

[16] For each variable, we calculated the 5th and 95th percentiles of the empirical distribution (P5 and P95, respectively) and trimmed observations outside the following range: P5 – 1 × (P95 – P5) to P95 + 1 × (P95 – P5). For normally distributed variables, this range covers approximately 4.95 standard deviations from the mean in each direction (= 1.65 + 1 × (1.65 − (−1.65)), which is more than 99.99% of the observations. For variables with relatively few outliers, the percentage of retained observations is also very high (often 100%). We repeated all the analyses using alternative outlier filters and estimation methodologies and confirmed the robustness of the findings. Also, our inferences endure if we winsorize instead of trim extreme values.

NCOs as a percentage of average gross loans is 0.53% (0.27%). In comparison, the means (medians) of *ExpectedRCL* and *RealizedRCL* are 0.50% (0.30%) and 0.55% (0.26%), respectively. The mean (median) loan yield in our sample is 6.73% (6.48%).

Turning to loan composition, real estate loans constitute about 68% of loans on average, with C&I loans a distant second at 17%. Consumer loans on average account for about 7% and all other loans combined constitute, on average, about 5%.[17]

Panel B of Table 1 reports the medians of the cross-sectional correlations over time between the variables used in our analyses. In general, *ExpectedRCL* is positively correlated with the other concurrent publicly disclosed credit risk measures. The Pearson correlation ranges from 0.95 with *RealizedRCL* to 0.46 with the ALLL; rank (Spearman) correlations are similar in magnitude. Also, *ExpectedRCL* is positively correlated with loan yield, confirming that banks charge higher interest on riskier loans to compensate for the expected credit losses.

V. MULTIVARIATE ANALYSIS

Estimating *ExpectedRCL*

To estimate *ExpectedRCL*, we perform quarterly cross-sectional regressions of Equation (9), using the trailing four quarters of data. Panel A of Table 2 presents the summary statistics from the 71 cross-sectional regressions (Q4:1997 to Q2:2015). For each estimated coefficient, we report the time-series mean of the coefficient, the time series t-statistic (the ratio of the time-series mean to the time-series standard error), and the time-series median of the cross-sectional t-statistic.

Most coefficients have the expected signs and are statistically significant. The most significant explanatory variable of the NCO rate is the one-period-lagged NCO rate, with a

[17] The variability of the proportion of "other loans" across the observations is small relative to that of the other loan categories, suggesting that the sum of the three explicit loan composition ratios—real estate, C&I, and consumer—has very low variability. Therefore, to mitigate multicollinearity, only two of these categories (i.e., RELoans and ConsLoans) are included in the regressions.

persistence parameter close to 0.5. Also highly significant are the unexpected change in NPLs (γ) and the level of NPLs (α_2). The γ coefficient is the proportion of a period's unexpected change in NPLs that represents a credit loss that has yet to be charged off. The estimated value of this parameter for the full sample period is approximately 0.17, implying that, on average, each dollar of unexpected NPLs is equivalent to 17 cents of credit loss not yet charged off. As we see in later analysis, this parameter differs across credit cycles in the expected direction. Loan yield and composition are also significantly associated with future credit losses. High-yield loans (α_3) and consumer loans (α_6) are, on average, riskier than other loans.

While Panel A of Table 2 summarizes the results of estimating Equation (9) over the full sample period, we expect the estimated coefficients to vary over time, especially with changes in the macro economy. We reflect this in two ways. First, in Panel B of Table 2, we present the time-series correlations between the estimated coefficients of Equation (9) and variables capturing the state of the macro economy. We use the percentage change in quarterly seasonally adjusted real gross domestic product (GDP) relative to the same quarter a year ago (%Δ in Real GDP) and the difference between Moody's seasoned Aaa and Baa corporate bond yields (Credit Spread) to proxy for economy-wide conditions. We find that the credit loss implications of the key credit-risk indicators included in Equation (9) are correlated with macroeconomic conditions in the expected direction. In particular, the proportion of the unexpected change in NPLs equivalent to credit losses yet to be charged off (γ) is positively (negatively) correlated with the credit spread (%Δ in Real GDP).

Second, in Figure 1, we plot the standardized coefficients and R-squared from the cross-sectional regressions of Equation (9) to illustrate their patterns over the sample period. The changes in coefficients around the financial crisis are particularly instructive. While the coefficients generally have the same signs in the crisis and non-crisis periods, the magnitudes of almost all of them changed significantly during the crisis. Both the persistence parameter (α_1) and the coefficient on NPLs (α_2) increased significantly. The proportion of the unexpected change in NPLs equivalent to credit losses yet to be charged off (γ) almost doubled, before returning to its pre-crisis level. Thus, credit losses since the beginning of the financial crisis increased not only because of the borrowers' deteriorating credit profiles, as reflected in NPLs and NCOs, but also because of greater loss implications of each dollar of unexpected change in NPLs and, possibly, because of more aggressive charge-off policies.

Evaluating the Predictive Ability of *ExpectedRCL*

The results presented in Table 2 suggest that the variables used to model *ExpectedRCL* are useful in explaining subsequent realized credit losses. However, these results do not directly provide evidence of the predictive ability of *ExpectedRCL*, which aggregates the information in the explanatory variables into a single measure of expected loss. To evaluate the predictive ability of *ExpectedRCL* for one-year-ahead realized credit losses and to compare it to the predictive ability of the other measures that reflect expected loss, we estimate cross-sectional regressions of the three models nested in the following specification:

$$NCO_{i,t+1} = \beta_{0,t} + \beta_{1,t}\widehat{ExpectedRCL}_{i,t} + \beta_{2,t}\frac{NCO_{i,t}}{AveLoans_{i,t}} + \varepsilon_{i,t+1} \qquad (11)$$

where $\widehat{ExpectedRCL}_{i,t}$ is estimated as described in Section III.[18] The results are reported in Table 3. Recall that part of our motivation is to identify a summary statistic that can indicate a bank's

[18] In this model, we measure realized credit losses using NCOs. An alternate measure is *RealizedRCL*. Since, as discussed above, *RealizedRCL* can remove some of the discretion in NCOs, we also investigate the predictive ability of *ExpectedRCL* using *RealizedRCL* as a measure of realized credit losses. Our inferences are unchanged. Therefore, for brevity and to avoid the concern that our results are an artifact of using a measure of realized credit losses constructed by us, we tabulate the results for which NCOs was used to measure realized credit losses.

one-year-ahead realized credit losses. Both current-year $\widehat{ExpectedRCL}$ (coefficient = 0.9640, median t-statistic = 19.6) and NCOs (coefficient = 0.7123, median t-statistic = 17.3) are positively associated with one-year-ahead NCOs when included in the model on a standalone basis. However, when both are included in the model together, only $\widehat{ExpectedRCL}$ is significant (coefficient = 0.9236, median t-statistic = 7.4). The coefficient on NCOs is positive but statistically insignificant (coefficient = 0.0131, median t-statistic = 0.8), suggesting that $\widehat{ExpectedRCL}$ reflects nearly all the information in the current year's NCOs relevant for one-year-ahead NCOs. Further, the mean estimated coefficient on $\widehat{ExpectedRCL}$ is close to 1, implying that, on average, each dollar of $\widehat{ExpectedRCL}$ translates into approximately a dollar of realized credit losses the next year.

Loans' Fair Value and *ExpectedRCL*

Given that the fair value of loans should capture information related to credit risk, interest rate risk, and other characteristics (e.g., Blankespoor et al. 2013; Cantrell et al. 2014), fair value measures can be expected to be related to future credit losses. Therefore, we investigate whether *ExpectedRCL* contains information relevant for the prediction of one-year-ahead realized credit losses incremental to that in the fair value of loans. We conduct these tests by estimating cross-sectional regressions of models that are nested in the following specifications:

$$\frac{NCO_{i,t+1}}{AveLoans_{i,t+1}} = \beta_{0,t} + \beta_{1,t}\frac{FVLoans_{i,t}}{Loans_{i,t}} + \beta_{2,t}\beta_{2,t}\widehat{ExpectedRCL}_{i,t} + \varepsilon_{i,t+1} \tag{12a}$$

$$\begin{aligned}\frac{NCO_{i,t+1}}{AveLoans_{i,t+1}} \\ = \beta_{0,t} + \beta_{1,t}\frac{\widehat{FVLoans\_Macro}_{i,t}}{Loans_{i,t}} + \beta_{2,t}\frac{\widehat{FVLoans\_Other}_{i,t}}{Loans_{i,t}} \\ + \beta_{3,t}\widehat{ExpectedRCL}_{i,t} \quad + \varepsilon_{i,t+1}\end{aligned} \tag{12b}$$

where FVLoans is the disclosed fair value of loans from the SNL Financial database. Since the fair value of loans reflects information beyond bank-specific credit losses (e.g., interest rates and economy-wide conditions), we also decompose FVLoans into a macroeconomic component (FVLoans_Macro) and a component that comprises all other information (FVLoans_Other). To do so, we estimate the following time-series bank-specific regressions:

$$\frac{FVLoans_{i,t}}{Loans_{i,t}} = \propto_{0,i} + \propto_{1,i} TBill_t + \propto_{2,i} TBond_t + \propto_{3,i} CSpread_t + \varepsilon_t \qquad (13)$$

Interest rate risk is an important component of the overall risk of loan portfolios and relates directly to changes in loan fair values. Hence, to extract the macroeconomic component of the fair value of loans, we regress FVLoans on the risk-free short- and long-term interest rate (e.g., Flannery and James 1984). TBill and TBond are the quarterly averages of the daily three-month Treasury bill secondary market rate and of the daily market yield on a 10-year US Treasury bond, respectively. To account for changes in the price of risk through the business cycle, we include the difference between Moody's seasoned Baa and Aaa corporate bond yields (CSpread). The predicted value (residual) of FVLoans from Equation (13) is $\mathrm{FVLo}\widehat{\mathrm{ans\_M}}\mathrm{acro}$ ($\mathrm{FVLo}\widehat{\mathrm{ans\_O}}\mathrm{ther}$). $\mathrm{FVLo}\widehat{\mathrm{ans\_M}}\mathrm{acro}$ reflects changes in the fair value of loans due to changes in overall market conditions. For example, banks highly sensitive to inverted term structures (i.e., TBond being less than TBill, a condition which predicts recessions) will likely have large credit losses during recessions. All other variables included in Equation (12) are as defined above.

US companies have been disclosing the fair value of most of their financial instruments—including loans—annually since 1992 and quarterly since the second quarter of 2009. SNL has collected this information since 2005. Our sample for this analysis includes 26 cross-sections (t): Q4:2005, Q4:2006, Q4:2007, Q4:2008, and Q2:2009–Q3:2014. We merge the fair value data with the FR Y-9C data using various identifiers and verify that the matches are correct. Panel A of

Table 4 presents the results of estimating the models nested in Equation (12a).[19] As expected, FVLoans is significantly related to one-year-ahead NCOs (coefficient = -0.0629, median t-statistic = -4.4). Importantly, $\widehat{ExpectedRCL}$ contains information, incremental to FVLoans, that is relevant to one-year-ahead NCOs. When $\widehat{ExpectedRCL}$ is included in the model with FVLoans, the coefficient on $\widehat{ExpectedRCL}$ is positive and significant (coefficient = 0.8105, median t-statistic = 12.6). Moreover, $\widehat{ExpectedRCL}$ subsumes the information in FVLoans relevant to one-year-ahead credit losses. When $\widehat{ExpectedRCL}$ is included in the model along with FVLoans, the coefficient on FVLoans is no longer statistically significant (median t-statistic = -1.1).

Panel B of Table 4 reports the results of the models nested in Equation (12b). When FVLoans is split into its two components, $\widehat{\text{FVLoans\_Macro}}$ is significantly associated with one-year-ahead NCOs (coefficient = -0.0729, median t-statistic = -3.7). However, the association between $\widehat{\text{FVLoans\_Other}}$ and subsequent NCOs is not significant (coefficient = -0.0245, median t-statistic = -0.79). When $\widehat{ExpectedRCL}$ is included in the model along with the two components of FVLoans, the coefficient on $\widehat{ExpectedRCL}$ is 0.8039 and highly significant (median t-statistic = 12.4). As with Panel A, when $\widehat{ExpectedRCL}$ is included in the model along with the components of FVLoans, the coefficients on $\widehat{\text{FVLoans\_Macro}}$ and $\widehat{\text{FVLoans\_Other}}$ are no longer statistically significant.

Arguably, the poor performance of FVLoans relative to *ExpectedRCL* in predicting the next year's realized credit losses could be due to the fact that fair values of loans are more

[19] The coefficients and the R-squared in Table 4 for the model that includes $\widehat{ExpectedRCL}$ as the only independent variable differ from those in Table 3 (and other tables) because, in Table 4, our sample is restricted to the observations for which fair value of loans is available.

informative for realized credit losses beyond the one-year horizon. We test this conjecture in an additional analysis in Section VI below.[20]

The ALLL, the PLLL, and *ExpectedRCL*

Next, we compare the overall and incremental information in *ExpectedRCL* about one-year-ahead realized credit losses relative to the ALLL and the PLLL. To this end, we estimate four cross-sectional regressions of models nested in the following specification:

$$\frac{NCO_{i,t+1}}{AveLoans_{i,t+1}} = \beta_{0,t} + \beta_{1,t}\widehat{ExpectedRCL}_{\iota,t} + \beta_{2,t}\frac{ALLL_{i,t}}{Loans_{i,t}} + \beta_{3,t}\frac{PLLL_{i,t}}{AveLoans_{i,t}} + \varepsilon_{i,t+1} \tag{14}$$

The results are reported in Table 5. The first result repeats the findings for $\widehat{ExpectedRCL}$ reported in Table 3. Next, we report the results of models that include the ALLL and the PLLL individually as explanatory variables. Both the ALLL (coefficient = 0.4127, median t-statistic = 7.9) and the PLLL (coefficient = 0.6717, median t-statistic = 18.1) are significantly associated, on a standalone basis, with one-year-ahead NCOs. While models with the PLLL and $\widehat{ExpectedRCL}$ as standalone explanatory variables have a mean R-squared above 40%, the model with the ALLL as the single explanatory variable has a mean R-squared of only 17%. When $\widehat{ExpectedRCL}$, ALLL, and PLLL are included together as explanatory variables, $\widehat{ExpectedRCL}$ continues to provide incremental information relevant to next year's NCOs (coefficient = 0.6046, median t-statistic = 7.2). Since the ALLL and the PLLL can reflect private information only available to

[20] Cantrell et al. (2014) find that the historical cost of loans better predicts NCOs (and NPLs) than the fair value of loans for both annual and aggregate multi-year NCOs. They conclude that this may result from the lack of scrutiny of the fair value measures. However, their multi-year results are not robust to including firm fixed effects (see their footnote 15).

managers, it is not surprising that they continue to provide incremental information. In summary, the evidence in Table 5 suggests that *ExpectedRCL* is incrementally useful beyond the ALLL and the PLLL in predicting one-year-ahead realized credit losses.

Out-of-sample Forecasting Ability of *ExpectedRCL*

Next, we investigate the out-of-sample forecasting performance of *ExpectedRCL* and other credit risk metrics for one-year-ahead NCOs. For *ExpectedRCL*, the forecast of one-year-ahead NCOs is the current period value of $\widehat{ExpectedRCL}$. For the other metrics, the forecasts are calculated using coefficient estimates from the following cross-sectional regressions:

$$\frac{NCO_{i,t}}{AveLoans_{i,t}} = \beta_{0,t} + \beta_{1,t} Credit\ Risk\ Metric_{i,t-1} + \varepsilon_{i,t} \qquad (15)$$

where Credit Risk Metric is either ALLL/Loans, PLLL/AveLoans, NCO/AveLoans, or FVLoans/Loans.[21] To assess the forecasting ability of each metric, we compare the absolute prediction errors based on $\widehat{ExpectedRCL}$ to the absolute prediction errors based on each of the other credit risk metrics (note that Equation (15) uses the same approach we used to derive $\widehat{ExpectedRCL}$, so the comparison puts all the metrics on an equal footing).

Table 6 reports distributional statistics of the absolute prediction errors for one-year-ahead NCOs using $\widehat{ExpectedRCL}$, ALLL, PLLL, NCO, and FVLoans. Panel A reports the results for our full sample. The mean (median) absolute prediction error for $\widehat{ExpectedRCL}$ is 0.293% (0.145%), which is lower than the mean (median) prediction error using any of the other credit risk metrics.[22] We also report the mean (median) of the difference between absolute prediction errors

[21] For example, to calculate the Q4:2010 forecast of one-year-ahead NCOs using the ALLL, we estimate Equation (15) in Q4:2010 where Credit Risk Metric is ALLL/Loans and apply the estimated coefficients to Q4:2010 values of the ALLL and Loans to obtain the forecasts.

[22] For the sample in which FVLoans is non-missing, the mean (median) of the absolute prediction error based on $\widehat{ExpectedRCL}$ is 0.354% (0.214%).

based on $\widehat{ExpectedRCL}$ and on the other credit risk metrics for each bank-quarter. The means and medians of the differences are negative and significant, indicating that $\widehat{ExpectedRCL}$ is a better predictor than the other metrics are of one-year-ahead NCOs. The improvement in the prediction of one-year-ahead NCOs offered by $\widehat{ExpectedRCL}$ is also economically significant. For example, relative to the ALLL, using $\widehat{ExpectedRCL}$ to predict one-year-ahead NCOs reduces the absolute prediction error for the average (median) bank by 24% (26%).[23]

Large banks tend to have more diversified loan portfolios and pursue riskier lending (Demsetz and Strahan 1997). Credit risk modeling to estimate loan loss accruals also differs across banks. Some depend on generic vendor-supplied models or simple spreadsheets to model the ALLL and the PLLL. Others develop customized multivariate statistical models using many characteristics, including underwriting criteria, current payment status, payment history, and relevant economic variables (Bhat et al. 2013). Since large banks have more resources and expertise, their ALLLs and PLLLs might better predict one-year-ahead credit losses. The discretion applied in the qualitative adjustments to the ALLL, the PLLL, and NCOs can also be expected to vary, although it is unclear how this would differ across banks of different sizes. Thus, it is plausible that other credit risk metrics are better than $\widehat{ExpectedRCL}$ at predicting one-year-ahead realized credit losses for subsamples of banks partitioned on size. Panels B and C report the results for such subsamples. For each quarter, banks with total assets greater than (equal to or below) the cross-sectional median are classified as large (small) banks. Consistent with our full sample results, the means and medians of the differences between the absolute prediction errors

[23] We calculate the economic significance for the average bank as the mean of the difference between absolute forecast errors based on *ExpectedRCL* and on ALLL/Loans (i.e., Mean Diff), scaled by the mean absolute prediction error for ALLL/Loans. The economic significance for the median bank is computed analogously.

based on $\widehat{ExpectedRCL}$ and on other credit risk metrics are negative and significant in both subsamples.

The methods used to estimate the ALLL and the PLLL vary across loan types. Generally, for homogeneous loans (e.g., residential real estate loans), loan loss reserves are estimated using statistical models based on past annualized loss experience. In contrast, large heterogeneous loans (e.g., commercial real estate and C&I loans) are evaluated individually for credit losses. For these loans, the recognition criteria for incurred credit losses are not met until shortly before default (Ryan 2007) and banks have more discretion in recognizing the losses (Berger and Udell 2002). Thus, for large heterogeneous loans, the ALLL and the PLLL may be less informative about one-year-ahead credit losses. We therefore explore the predictive ability of $ExpectedRCL$ for subsamples based on the concentration of real-estate-related loans and of C&I loans.

Panel D (E) reports the results for banks whose holdings of commercial (residential) real estate loans as a proportion of total loans is above the cross-sectional median.[24] We continue to find the predictive ability of $\widehat{ExpectedRCL}$ for one-year-ahead NCOs is superior to that of ALLL, PLLL, NCO, and FVLoans. The mean and median of the differences between absolute prediction errors based on $\widehat{ExpectedRCL}$ and on other credit risk metrics are negative and significant.

Finally, we partition the sample based on whether the holdings of C&I loans as a proportion of total loans are above (equal to or below) the cross-sectional median. Panel F (G) reports the results for the subsample with high (low) concentration of C&I loans. $\widehat{ExpectedRCL}$ continues to have better predictability than the other credit risk metrics for one-year-ahead NCOs in both

[24] Commercial real estate loans are real estate loans secured by nonfarm nonresidential properties. All other loans secured by real estate are classified as residential real estate loans.

subsamples. The mean and median of the difference in absolute prediction errors based on $\widehat{ExpectedRCL}$ relative to those based on other credit risk metrics is negative and significant.

In conclusion, the findings of our out-of-sample tests indicate that *ExpectedRCL* has better predictive ability for one-year-ahead NCOs than ALLL, PLLL, NCO, and FVLoans in the full sample and in subsamples based on size and on loan portfolio concentration.

VI. ADDITIONAL ANALYSES AND ROBUSTNESS TESTS

We conduct several additional analyses to establish the usefulness of *ExpectedRCL* and test its robustness as a summary indicator of one-year-ahead realized credit losses.

Comparison of *ExpectedRCL* and Analyst Forecasts of the PLLL

One motivation for developing a summary measure of expected credit losses is that analysts and investors can use it to forecast earnings. We evaluate the usefulness of *ExpectedRCL* to investors, analysts, and other users by comparing the ability of *ExpectedRCL* and of analyst PLLL forecasts to predict the PLLL. The SNL Financial database provides analyst forecasts of the PLLL beginning in the third quarter of 2008. Since *ExpectedRCL* is an estimate of one-year-ahead realized credit losses, we restrict our analysis to bank-quarters with available annual forecasts of the PLLL. During our sample period, we identified 1,344 such bank-quarters representing 319 unique banks.

Table 7 reports the distribution statistics of absolute prediction errors of forecasting the PLLL using $\widehat{ExpectedRCL}$ and using the mean analysts' PLLL forecast (consensus PLLL forecast). The prediction errors are scaled by total assets and expressed as a percentage.[25] We estimate the prediction errors using the most recent forecasts that were available before the earnings announcement. In our full sample, the mean (median) absolute prediction error using

[25] Our inferences are unchanged if we use the median of analysts' PLLL forecasts instead.

$\widehat{ExpectedRCL}$ is 0.208% (0.137%), which is smaller than the mean (median) absolute prediction error using the consensus PLLL forecast, 0.423% (0.192%). The mean and median of the difference between the absolute prediction errors based on $\widehat{ExpectedRCL}$ and on the consensus PLLL forecast is negative and significant, suggesting that $\widehat{ExpectedRCL}$ predicts the PLLL better than analyst PLLL forecasts do. This improvement in predicting the PLLL is economically meaningful. For example, in the full sample, using $\widehat{ExpectedRCL}$ to forecast the PLLL rather than using the consensus PLLL forecast reduces the absolute prediction error by 51% (30%) for the average (median) bank. We find similar results in subsamples partitioned on bank size and on loan portfolio concentration.

Comparison of Conditional Earnings Surprises

Earnings surprises have been studied extensively in the literature and have important implications in practice. As an additional test of the potential usefulness of $\widehat{ExpectedRCL}$, we compare the earnings surprise conditional on the absolute difference between $\widehat{ExpectedRCL}$ and (a) the consensus PLLL forecasts and (b) forecasts of one-year-ahead NCOs based on NCO/AveLoans, ALLL/Loans, or PLLL/AveLoans, computed using Equation (15). Earnings surprise is calculated as the absolute difference between reported net income and the mean of analysts' net income forecasts (consensus earnings forecast) divided by total assets, the ratio being expressed as a percentage.[26]

Table 8 reports the results. When the absolute difference between $\widehat{ExpectedRCL}$ and the consensus PLLL forecast is less than or equal to the cross-sectional median, the mean (median) earnings surprise is 0.289% (0.150%). In comparison, when the difference between $\widehat{ExpectedRCL}$

[26] Our inferences are unchanged if we use the median of analysts' net income forecasts instead.

and consensus PLLL forecast is greater than the cross-sectional median, the mean (median) earnings surprise is 0.836% (0.409%). The differences in the mean and median earnings surprises between the two partitions are significant, indicating that banks, on average, experience larger earnings surprises when the differences between $\widehat{ExpectedRCL}$ and analyst PLLL forecasts are larger. Our inferences using forecasts of one-year-ahead NCOs are similar. For example, when the absolute difference between $\widehat{ExpectedRCL}$ and a forecast of one-year-ahead NCOs based on NCO/AveLoans is less than or equal to the cross-sectional median, the mean (median) earnings surprise is 0.442% (0.191%). In comparison, the mean (median) earnings surprise is 0.686% (0.292%) when the absolute difference is above the cross-sectional median. Also, the differences in the mean and median earnings surprises between the two partitions are significant.

Using *ExpectedRCL* to Predict Bank Failure

As an additional test of the potential usefulness of $ExpectedRCL$, we consider its ability to predict the extreme outcome of credit losses; namely, bank failure. Using the Federal Deposit Insurance Corporation website, we identify 404 banks that failed during our sample period.[27] We drop to 184 banks when we match each failed bank to its parent bank holding company and end up with 71 failed banks in our sample after imposing the data requirements. We estimate logit regressions of models nested in the following specification:

[27] https://www.fdic.gov/bank/individual/failed/banklist.html

$$
\begin{aligned}
\Pr(Fail = 1)_t = {} & \beta_0 + \beta_1 \frac{Equity_t}{Assets_t} + \beta_2 \frac{Loans_t}{Assets_t} + \beta_3 \frac{RELoans_t}{Loans_t} + \beta_4 \frac{C\&ILoans_t}{Loans_t} \\
& + \beta_5 \frac{OREO_t}{Assets_t} + \beta_6 \frac{EarnNC_t}{Assets_t} + \beta_7 \frac{NPL_t}{Assets_t} + \beta_8 ROE_t + \beta_9 LIQ_t \\
& + \beta_{10} SIZE_t + \beta_{11} AGE_t + \beta_{12} \widehat{ExpectedRCL}_t + \beta_{13} \frac{ALLL_t}{Loans_t} \\
& + \beta_{14} \frac{PLLL_t}{AveLoans_t} + \beta_{15} \frac{NCO_t}{AveLoans_t} + \varepsilon_t
\end{aligned} \tag{16}
$$

where Fail is an indicator variable that equals 1 if a bank fails within one year of the end of quarter t. We restrict our horizon of bank failure to one year as $\widehat{ExpectedRCL}$ is an estimate of one-year-ahead realized credit losses. We selected the explanatory variables included in the model based on measures used in the regulatory evaluation of bank health and in prior studies of bank failure (e.g., Thomson 1991; Wheelock and Wilson 2000). Equity is a bank's total equity. OREO is other real estate owned by a bank. EarnNC is total income earned but not collected. ROE is return on equity. LIQ is net federal funds purchased scaled by total assets. SIZE is the log of total assets and AGE is the log of the bank's age in years. The model also includes $\widehat{ExpectedRCL}$ and three other credit risk metrics—the ALLL, the PLLL, and NCO. We do not include FVLoans because of the sparse availability of this data.[28] All other variables are as previously defined. We cluster standard errors by bank.

The results of our bank failure tests should be considered with caution, as the number of failing banks in the sample is relatively small. The results of estimating Equation (16) using logit regressions are reported in Panel A of Table 9. In Column 1, we estimate a model that excludes $\widehat{ExpectedRCL}$ as a predictor of bank failure. The coefficient on PLLL is 153.85 and significant

[28] If FVLoans is included in the model, only nine failed banks remain in the sample.

(p-value < 0.01), suggesting that the PLLL is useful in predicting failure. The coefficient on ALLL is negative and marginally significant (p-value < 0.10), suggesting that, in spite of the relatively large PLLL immediately prior to failure, the ALLL remains understated. Interestingly, the coefficient on NCOs is negative and significant (coefficient = -168.01, p-value < 0.010). We interpret the negative coefficient to suggest that, controlling for the PLLL and the ALLL, healthier banks are timelier in charging off loans.[29]

In Column 2, we estimate Equation (16), including $\widehat{ExpectedRCL}$ as an independent variable. After controlling for other credit risk metrics and bank characteristics, $\widehat{ExpectedRCL}$ is incrementally useful in predicting bank failure over the next year. The coefficient on $\widehat{ExpectedRCL}$ is 139.51 and significant (p-value < 0.01). The coefficient on PLLL drops to 94.87 but remains significant (p-value < 0.01), indicating that the PLLL contains incremental information relevant for predicting bank failure beyond $\widehat{ExpectedRCL}$ and the other variables included in the model. The coefficient on ALLL becomes insignificant, suggesting that $\widehat{ExpectedRCL}$ subsumes all the information in the ALLL relevant to future bank failure. The coefficient on *NCO* continues to be significant (p-value < 0.01), but negative.[30]

As a robustness test, we also estimate a Cox (1972) proportional-hazard model of bank failure to investigate the usefulness of $\widehat{ExpectedRCL}$ for predicting bank failure. The results are reported in Panel B of Table 9. Our inferences remain unchanged. We continue to find that, after controlling for other credit risk metrics and bank characteristics, a higher $\widehat{ExpectedRCL}$ is

[29] Liu and Ryan (2006) report that during the 1990s, banks accelerated provisioning for loan losses to smooth earnings and accelerated loan charge-offs to mask the smoothing. These effects were stronger for more profitable banks.

[30] In untabulated analyses, we compute absolute prediction errors of bank failure using the two logit regression models reported in Panel A of Table 9. While the mean absolute prediction error is smaller when $\widehat{ExpectedRCL}$ is included in the model as a predictor, the difference between the means is insignificant, which is partially attributable to low power due to the very low occurrence of bank failure during our sample period.

associated with an increase in the failure hazard rate of banks. The coefficient on $\widehat{ExpectedRCL}$ is 98.95 and significant (p-value < 0.01).

Prediction of Long-horizon Net Charge-offs

So far, we have investigated the predictive ability of *ExpectedRCL* only for the next year's realized credit losses. We have done so because one-year-ahead prediction of credit losses is consistent with the period used in the estimation of the ALLL under certain regulatory guidance, because it is considered in the assessment of annual earnings estimates, and because it is the objective of IFRS 9 for most loans. However, we also test for the possibility that *ExpectedRCL* provides useful information for credit losses beyond one year.

Table 10 reports the results from regressions that compare the predictive ability of *ExpectedRCL* to that of other credit risk measures for NCOs aggregated over the next three years. Specifically, we estimate regression models nested in the following specification:

$$\frac{\sum_{i=1}^{3} NCO_{i,t+i}}{\sum_{i=1}^{3} AveLoans_{i,t+i_t}}$$

$$= \beta_{0,t} + \beta_{1,t}\widehat{ExpectedRCL}_{i,t} + \beta_{2,t}\frac{ALLL_{i,t}}{Loans_{i,t}} + \beta_{3,t}\frac{PLLL_{i,t}}{AveLoans_{i,t}} \tag{17}$$

$$+ \beta_{4,t}\frac{NCO_{i,t}}{AveLoans_{i,t}} + \beta_{5,t}\frac{FVLoans_{i,t}}{Loans_{i,t}} + \varepsilon_{i,t+1}$$

We find that *ExpectedRCL* is economically and statistically significant in predicting NCOs over the next three years in standalone and multivariate regressions that include current-period ALLL, PLLL, NCOs, and FVLoans. In the model in which $\widehat{ExpectedRCL}$ is included as an explanatory variable along with ALLL, PLLL, and NCOs (ALLL, PLLL, NCOs, and FVLoans), the mean coefficient on $\widehat{ExpectedRCL}$ is 0.8014 (0.5546) with a median t-statistic of 5.2 (5.0). Thus, *ExpectedRCL* appears to be useful for improving the prediction not only of one-year-ahead credit losses but also of longer-term credit losses.[31] Also, the fair value of loans does not contain information incremental to *ExpectedRCL* and other credit-risk metrics relevant for long-horizon net charge-offs. The coefficient of -0.0034 on FVLoans is insignificant (median t-statistic = -0.26).

Robustness Tests

In our main analyses, we use only the most recent coefficients of Equation (9) to estimate *ExpectedRCL*, which assumes that there is no incremental information in prior coefficient estimates. Given the limited size of each cross-section and the substantial variability of unexpected credit losses, we may be relinquishing statistical power by excluding the longer period of data in estimating the coefficients. On the other hand, as we show in Figure 1, the coefficients change considerably over time, so the most recent estimates are likely to be less biased than prior estimates. To evaluate this bias/noise trade-off, we repeat the analysis, extrapolating from the time-series of the coefficient estimates. We replicate the first regression model of Table 3 using five alternative estimates of *ExpectedRCL*, derived using (a) a moving average of the last four coefficient estimates and (b) exponential smoothing of the coefficient estimates, with a smoothing factor of 0.5, 0.25, 0.1, and 0.01.[32] In untabulated results, we find that extrapolating from past coefficient estimates improves the accuracy of credit loss forecasts slightly, particularly when using low exponential smoothing factors (i.e., when giving higher weight to past estimates).

Next, since future (i.e., t+1) net charge-offs and the unexpected change in NPLs used in estimating the parameters of Equations (7) and (9) are measured using the trailing four quarters of data, the explanatory variables of Equation (9) are at least a year old at the time the parameters are

[31] In untabulated analyses, we examine the predictive ability of *ExpectedRCL* using two- and three-year-ahead NCOs as measures of future realized credit losses rather than NCOs aggregated over three years. We continue to find that *ExpectedRCL* provides information incremental to ALLL, PLLL, NCOs, and FVLoans relevant for the prediction of two- and three-year-ahead realized credit losses.

[32] The predicted value (s_{t+1}) for the series x_t under exponential smoothing with a smoothing factor f is $f*x_t + (1-f)*s_t$, where $s_1=x_0$.

estimated. An alternative approach, which allows for a substantially shorter delay, is to measure net charge-offs and the unexpected change in NPLs using annualized quarterly data. This allows the estimated parameters to reflect more recent information but at the cost of using seasonal and potentially noisy information. We replicate the first regression model of Table 3 with *ExpectedRCL* estimated using annualized credit losses; untabulated results suggest that this slightly improves its predictive ability.

VII. CONCLUSION

Estimating one-year-ahead expected credit losses has vexed analysts and investors and has recently drawn renewed interest from regulators with the IASB's passage of IFRS 9. Our study develops a metric of one-year-ahead expected credit losses using a linear combination of several credit-risk-related measures disclosed by banks.

We develop a structural model of the expected rate of credit losses (*ExpectedRCL*) and estimate this measure using time-varying coefficients from cross-sectional regressions and bank-specific periodic disclosures. The resulting empirical measure of *ExpectedRCL* substantially outperforms the historical rate of net charge-offs, the ALLL, the PLLL, and FVLoans in predicting one-year-ahead realized credit losses. This result for the full sample is robust when we partition banks on size and on loan type.

We also find that *ExpectedRCL* has, on average, better predictive ability for the one-year-ahead PLLL than analysts' forecasts of the PLLL. Further, banks have larger earnings surprises relative to analysts' estimates when the difference between *ExpectedRCL* and the analyst's estimate of the PLLL or between *ExpectedRCL* and forecasts of one-year-ahead realized credit losses based on the ALLL, the PLLL, or NCOs is larger. Finally, we find that *ExpectedRCL* is incrementally useful in predicting bank failures over the next year.

An extension for future research relates to the potential use of *ExpectedRCL* in the Capital and Loss Assessment under Stress Scenarios (CLASS) model, which has come to be considered increasingly useful for conducting top-down stress tests using only publicly available data (Hirtle et al. 2015). In projecting loan-related expenses and credit losses, the CLASS model uses time-series models to project the NCO rate as a function of the lagged NCO rate and macroeconomic variables that are adjusted to reflect the different stress scenarios. Evidence from in- and out-of-sample tests presented above suggests that *ExpectedRCL* is a better predictor of the next period's NCO rate than lagged NCOs is. Also, *ExpectedRCL* subsumes nearly all the information in NCOs relevant for predicting credit losses. Thus, we believe *ExpectedRCL* can be useful in CLASS models to improve projections of the expected credit losses.

APPENDIX A: VARIABLE DEFINITIONS

This table provides detailed definitions of the variables used in our analyses. The data source and variable construction is described in brackets.

%Δ in Real GDP	Percentage change in quarterly, seasonally adjusted real gross domestic product. [Federal Reserve Economic Data: GDPC1]
Age	Log of bank age in years. Age is measured beginning the first time a bank appears in the FR-Y9C database.
ALLL	Allowance for loan and lease losses. [FR-Y9C: BHCK3123]
Assets	Total assets. [FR-Y9C: BHCK2170]
AveLoans	Average of the beginning and ending amounts of loans and leases held for investment for the quarter.
Charge-offs	Loan charge-offs. [FR-Y9C: BHCK4635]
Credit Spread	Difference between Moody's seasoned Aaa and Baa corporate bond yields. [Federal Reserve Economic Data: Aaa – Baa]
ConsLoans	Loans categorized as consumer loans. Typically comprises loans provided to individuals for household, family, and other personal expenditures. [FR-Y9C: BHCKB538 + BHCKB539 + BHCKK137 + BHCKK207 + BHCK2011 + BHCK2008]
C&I Loans	Commercial and industrial loans. [FR-Y9C: BHCK1763 + BHCK1764]
Commercial Real Estate Loans	Real estate loans secured by nonfarm nonresidential properties. [FR-Y9C: BHDM1480 till Q4:2006; BHCK2122+BHCK2123 since Q1:2007]
CSpread	Difference between Moody's seasoned Baa and Aaa corporate bond yields. [Federal Reserve Economic Data: Baa – Aaa]
Delinquent Loans	Total loans past due 90 days or more and still accruing interest. Delinquent loans guaranteed or otherwise protected by the US government are excluded. [FR-Y9C: BHCK5525 – BHCK3506 – BHCKK040 – BHCKK043 – BHCKK103 – BHCK5616 – BHCKC867]
EarnNC	Income earned but not collected. [FR-Y9C: BHCK5397 till Q4:2000 and BHCKB556 since Q1:2001]
Equity	Total equity [FR-Y9C: BHCK310]
ExpectedRCL	Expected rate of credit losses as defined in Equation (10). For each quarter, Equation (9) is estimated using the trailing four quarter of data to derive coefficients used in Equation (10).
FloatLoanRatio	Proportion of loans that reprice or mature within one year. [FR-Y9C: (BHCK3197 – BHCK0395 – BHCK0397 – Federal Funds Sold and Securities Purchased under Agreements to Resell)/(BHCK2122 + BHCK1754 + BHCK1773 – BHCK5526). Federal Funds Sold and Securities Purchased under Agreements to Resell is calculated as BHCK0276 + BHCK0277 till Q4:1996; BHCK1350 between Q1:1997 and Q4:2001; and BHDMB987 + BHCKB989 since Q2:2001]
FVLoans	Fair value of net loans. [SNL Financial: FV_NET_LOAN]
FVLoans_Macro	Macroeconomic component of FVLoans calculated as the predicted value of FVLoans from firm-specific regressions of Equation (13).
FVLoans_Other	The non-macroeconomic component of FVLoans calculated as the residual from firm-specific regressions of Equation (13).
LIQ	Federal funds purchased (FFP) minus federal funds sold (FFS), scaled by Assets. [FR-Y9C: FFP = BHCK0278+BHCK0279 till Q4:1996; BHCK2800 between Q1:1997 and Q4:2000; BHDMB993+BHCKB995 since Q1:2001; FFS = BHCK0276+BHCK0277 till Q4:1996; BHCK1350 between Q1:1997 and Q4:2000; BHDMB987+BHCKB989 since Q12001]

Loans	Total loans and leases held for investment. [FR-Y9C: BHCK2122 – BHCK5369]
LoansYield	Annualized tax equivalent interest rate on loans. [FR-Y9C: 4 x (Tax Equivalent Interest on Loans/Quarterly average of Loans and Leases). Interest on Loans is calculated as BHCK4393 + BHCK4503 + BHCK4504 + BHCK4059 + BHCK4505 + BHCK4307 till Q4:2000; BHCK4010 + BHCK4059 + BHCK4065 between Q1:2001 and Q4:2007; and BHCK4435 + BHCK4436 + BHCKF821 + BHCK4059 + BHCK4065 since Q1:2008. The tax adjustment is calculated as (BHCK4504 + bhck4307) x (0.35/(1-0.35)) till Q4:2000; and BHCK4313 x (0.35/(1 – 0.35)) since Q1:2001]
Nonaccruing Loans	Total loans not accruing interest, excluding those guaranteed or protected by the US government. [FR-Y9C: BHCK5526 – BHCK3507 – BHCKK041 – BHCKK044 – BHCKK104 – BHCK5617 – BHCKC868]
NCO	Net charge-offs. [FR-Y9C: BHCK4635 – BHCK4605]
NPL	Nonperforming loans. [Delinquent Loans + Nonaccruing Loans + Restructured Loans]
ΔNPL^{unexp}	Unexpected change in nonperforming loans as defined in Equation (4).
OREO	Other real estate owned. [FR-Y9C: BHCK2150; if BHCK2150 is missing, then BHCK2744 + BHCK2745]
OthrLoans	Total of all loans that are not categorized as consumer, commercial and industrial, or real estate loans. [Loans – RELoans – C&I Loans – ConsLoans]
PLLL	Provision for loan and lease losses. [FR-Y9C: BHCK4230 + BHCK4243]
RealizedRCL	Realized rate of credit losses as defined in Equation (5). For each quarter, Equation (9) is estimated using the trailing four quarters of data to derive γ_t used in Equation (5) to calculate *RealizedRCL*.
RELoans	Loans secured by real estate. [FR-Y9C: BHCK1410]
Residential Real Estate Loans	RELoans minus Commercial Real Estate Loans.
Restructured Loans	Total loans restructured during the quarter. Typically, restructuring of loans involves a reduction of either interest or principal because of deterioration in the borrower's financial position. [FR-Y9C: BHDMK158 + BHDMK159 + BHDMF576 + BHDMK160 + BHDMK161 + BHDMK162 + BHCKK163 + BHCKK164 + BHCKK165 + BHCK1616]
ROE	Return on equity, calculated as net income available to common shareholders scaled by average equity. [FR-Y9C: (BHCK4340-BHCK4598)/Average Equity]
SIZE	Log of Assets.

REFERENCES

Ahmed, A., C. Takeda, and S. Thomas. 1999. Bank loan loss provisions: A reexamination of capital management, earnings management, and signaling effects. *Journal of Accounting and Economics* 28 (1): 1–25.

Araten, M., M. Jacobs, Jr., and P. Varshney. 2004. Measuring LGD on commercial loans: An 18-year internal study. *The RMA Journal* (May): 92–97.

Barth, M., W. Beaver, and W. Landsman. 1996. Value-relevance of banks' fair value disclosures under SFAS No. 107. *The Accounting Review* 71 (4): 513–537.

Beatty, A., S. Chamberlain, and J. Magliolo. 1995. Managing financial reports of commercial banks: The influence of taxes, regulatory capital, and earnings. *Journal of Accounting Research* 33 (2): 231–261.

Beaver, W. H., C. Eger, S. Ryan, and M. Wolfson. 1989. Financial reporting, supplemental disclosures, and bank share prices. *Journal of Accounting Research* 27 (2): 157–178.

Beaver, W. H., and E. E. Engel. 1996. Discretionary behavior with respect to allowance for loan losses and the behavior of security prices. *Journal of Accounting and Economics* 22 (1-3): 177–206.

Berger, A., and G. Udell. 2002. Small business credit availability and relationship lending: The importance of bank organizational structure. *The Economic Journal* 122: 32–53.

Bhat, G., S. G. Ryan, and D. Vyas. 2013. The implications of banks' credit risk modeling disclosures for their loan loss provision timeliness and loan origination procyclicality. Working paper, New York University.

Blankespoor, E., T. Linsmeier, K. Petroni, and C. Shakespeare. 2013. Fair value accounting for financial instruments: Does it improve the association between bank leverage and credit risk? *The Accounting Review* 88 (4): 1143–1177.

Bushman, R. M., and C. D. Williams. 2015. Delayed expected loss recognition and the risk profile of banks. *Journal of Accounting Research* 53 (3): 511-553.

Calomiris, C., and D. Nissim. 2014. Crisis-related shifts in the market valuation of banking activities. *Journal of Financial Intermediation* 23 (3): 400–435.

Cantrell, B., J. McInnis, and C. Yust. 2014. Predicting credit losses: Loan fair values versus historical costs. *The Accounting Review* 89 (1): 147–176.

Collins, J. H., D. A. Shackelford, and J. M. Wahlen. 1995. Bank differences in the coordination of regulatory capital, earnings, and taxes. *Journal of Accounting Research* 33 (2): 263–291.

Cox, D. R. 1972. Regression models and life tables (with discussion). *Journal of the Royal Statistical Society*, Series B 34: 187-220.

Demsetz, R. S., and P. E. Strahan. 1997. Diversification, size, and risk at bank holding companies. *Journal of Money, Credit and Banking* 29 (3): 300–313.

Elliott, J. A., J. D. Hanna, and W. H. Shaw. 1991. The evaluation by the financial markets of changes in bank loan loss reserve levels. *The Accounting Review* 66 (4): 847–861.

Ernst and Young. 2014. IASB Projects: A Pocketbook Guide. Available at: http://www.ey.com/Publication/vwLUAssets/EY-iasb-projects-a-pocketbook-guide-as-at-31-march-14/$FILE/EY-PBGuide-April-2014.pdf

Federal Deposit Insurance Corporation. 2006. Interagency Policy Statement on the Allowance for Loan and Lease Losses. Available at: https://www.fdic.gov/news/news/financial/2006/fil06105a.pdf.

Flannery, M. J., and C. M. James. 1984. The effect of interest rate changes on the common stock returns of financial institutions. *The Journal of Finance* 39 (4): 1141-1153.

Griffin, P. A., and S. J. R. Wallach. 1991. Latin American lending by major U.S. banks: The effects of disclosures about nonaccrual loans and loan loss provisions. *The Accounting Review* 66 (4): 830–846.

Hirtle, B., A. Kovner, J. Vickery, and M. Bhanot. 2015. Assessing Financial Stability: The Capital and Loss Assessment under Stress Scenarios (CLASS) Model. Federal Reserve Bank of New York Staff Report 663, July.

Khan, U., and N. B. Ozel. 2016. Real activity forecasts using loan portfolio information. *Journal of Accounting Research* 54 (3): 895–937.

Liu, C., and S. G. Ryan. 2006. Income smoothing over the business cycle: Changes in banks' coordinated management of provisions for loan losses and loan charge-offs from the pre-1990 bust to the 1990s boom. *The Accounting Review* 81 (2): 421–441.

Liu, C., S. G. Ryan, and J. Wahlen. 1997. Differential valuation implications of loan loss provisions across banks and fiscal quarters. *The Accounting Review* 72 (1): 133–146.

Moyer, S. E. 1990. Capital adequacy ratio regulations and accounting choices in commercial banks. *Journal of Accounting and Economics* 13 (2): 123–154.

Nissim, D. 2003. Reliability of banks' fair value disclosure for loans. *Review of Quantitative Finance and Accounting* 20 (4): 355–384.

Ryan, S. 2007. Financial instruments and institutions: Accounting and disclosure rules. Second edition. New York: John Wiley & Sons.

Thomson, J. B. 1991. Predicting bank failures in the 1980s. *Federal Reserve Bank of Cleveland Economic Review* 27: 3–15.

Vyas, D. 2011. The timeliness of accounting write-downs by U.S. financial institutions during the financial crisis of 2007–2008. *Journal of Accounting Research* 49 (3): 823–860.

Wahlen, J. M. 1994. The nature of information in commercial bank loan loss disclosures. *The Accounting Review* 69 (3): 455–478.

Wheelock, D. C., and P. W. Wilson. 2000. Why do banks disappear? The determinants of U.S. bank failures and acquisitions. *Review of Economics and Statistics* 82 (1): 127-138.

TABLE 1
Summary Statistics

Panel A: Distribution Statistics

		Obs.	Mean	SD	Min	Q1	Med.	Q3	Max
V1	*NCO/AveLoans*	33,929	0.53%	0.76%	-2.06%	0.11%	0.27%	0.60%	4.84%
V2	ΔNPL^{Unexp} */AveLoans*	33,644	0.08%	1.14%	-5.35%	-0.33%	-0.02%	0.33%	6.23%
V3	*RealizedRCL*	33,362	0.55%	0.84%	-2.51%	0.09%	0.26%	0.62%	7.21%
V4	*NPL/Loans*	34,845	1.77%	2.08%	0.00%	0.51%	1.01%	2.18%	13.44%
V5	*RELoans /Loans*	35,306	68.45%	19.90%	0.00%	58.68%	71.86%	82.78%	100.00%
V6	*ConsLoans /Loans*	34,878	7.41%	8.80%	0.00%	1.34%	3.91%	10.63%	57.30%
V7	*C&Iloans /Loans*	35,224	17.02%	11.74%	0.00%	9.15%	14.94%	22.07%	77.31%
V8	*OthrLoans /Loans*	34,642	4.94%	5.94%	0.00%	0.90%	2.95%	6.69%	39.84%
V9	*LoansYield*	34,181	6.73%	1.72%	-0.14%	5.50%	6.48%	7.91%	15.19%
V10	*FloatLoanRatio*	35,320	39.66%	17.89%	0.00%	26.35%	39.11%	51.39%	100.00%
V11	*ExpectedRCL*	32,848	0.50%	0.67%	-1.28%	0.16%	0.30%	0.57%	6.74%
V12	*ALLL /Loans*	35,026	1.58%	0.76%	0.00%	1.14%	1.41%	1.81%	5.71%
V13	*PLLL/AveLoans*	33,904	0.62%	0.84%	-2.51%	0.17%	0.35%	0.71%	5.46%
V14	*FVLoans /Loans*	7,063	0.984	0.032	0.808	0.974	0.988	1.000	1.130

The sample period is Q4:1996 through Q2:2015. The disclosed fair value of loans is available for a subset of firms in 26 quarters (Q4:2005, Q4:2006, Q4:2007, Q4:2008, and Q2:2009–Q3:2014). Balance sheet items are measured at the end of the quarter. Income statement items are measured using the trailing four quarters of data. Details on variable definitions are provided in Section III and Appendix A.

Panel B: Medians (over Time) of Cross-sectional Pearson (below the Diagonal) and Spearman (above) Correlation Coefficients

		V1	V2	V3	V4	V5	V6	V7	V8	V9	V10	V11	V12	V13	V14
V1	*NCO/AveLoans*	.	0.03	0.93	0.52	-0.37	0.32	0.18	0.21	0.14	0.15	0.85	0.35	0.77	-0.46
V2	*ΔNPL^{unexp} /AveLoans*	0.02	.	0.25	0.19	0.00	0.00	-0.03	0.00	0.03	0.03	0.22	0.00	0.14	0.08
V3	*RealizedRCL*	0.97	0.23	.	0.55	-0.32	0.31	0.16	0.18	0.12	0.16	0.91	0.33	0.76	-0.38
V4	*NPL/Loans*	0.43	0.22	0.46	.	0.01	-0.01	0.06	0.02	0.16	0.08	0.67	0.40	0.42	-0.31
V5	*RELoans /Loans*	-0.27	0.03	-0.24	0.10	.	-0.50	-0.67	-0.58	0.05	-0.23	-0.41	-0.16	-0.15	0.02
V6	*ConsLoans /Loans*	0.30	-0.01	0.29	-0.06	-0.46	.	0.10	0.32	0.06	-0.08	0.33	0.06	0.19	-0.05
V7	*C&Iloans /Loans*	0.01	-0.01	0.01	-0.07	-0.59	-0.02	.	0.35	-0.06	0.30	0.19	0.20	0.11	-0.14
V8	*OthrLoans /Loans*	0.06	0.00	0.07	-0.03	-0.53	0.13	0.22	.	0.00	0.11	0.23	0.14	0.06	0.00
V9	*LoansYield*	0.23	0.03	0.24	0.15	0.07	0.08	-0.08	-0.04	.	-0.08	0.24	0.16	0.22	0.08
V10	*FloatLoanRatio*	0.09	0.02	0.10	0.05	-0.30	-0.10	0.30	0.09	-0.06	.	0.25	0.16	0.17	-0.20
V11	*ExpectedRCL*	0.90	0.27	0.95	0.67	-0.32	0.29	0.07	0.14	0.33	0.20	.	0.38	0.72	-0.38
V12	*ALLL /Loans*	0.54	0.01	0.52	0.41	-0.07	0.06	0.14	0.05	0.25	0.09	0.46	.	0.34	-0.53
V13	*PLLL/AveLoans*	0.89	0.12	0.89	0.40	-0.22	0.24	-0.01	0.04	0.35	0.10	0.82	0.52	.	-0.29
V14	*FVLoans /Loans*	-0.36	0.07	-0.32	-0.26	0.02	-0.03	-0.06	0.02	0.14	-0.16	-0.33	-0.44	-0.22	.

The sample period is Q4:1996 through Q2:2015. The disclosed fair value of loans is available for a subset of firms in 26 quarters (Q4:2005, Q4:2006, Q4:2007, Q4:2008, and Q2:2009–Q3:2014). Balance sheet items are measured at the end of the quarter. Income statement items are measured using the trailing four quarters of data. Details on variable definitions are provided in Section III and Appendix A.

TABLE 2
Summary Statistics from Cross-sectional Regressions for Estimating *ExpectedRCL*

$$\frac{NCO_{i,t}}{AveLoans_{i,t}} = \propto_{0,t}{}^{*} + \propto_{1,t} \frac{NCO_{i,t-1}}{AveLoans_{i,t-1}} + \propto_{1,t} \gamma_t \frac{\Delta NPL_{i,t-1}^{unexp}}{AveLoans_{i,t-1}} + \propto_{2,t} \frac{NPL_{i,t-1}}{Loans_{i,t-1}} + \propto_{3,t} LoansYield_{i,t-1}$$
$$+ \propto_{4,t} FloatLoanRatio_{i,t-1} + \propto_{5,t} \frac{RELoans_{i,t-1}}{Loans_{i,t-1}} + \propto_{6,t} \frac{ConsLoans_{i,t-1}}{Loans_{i,t-1}} + \varepsilon_{i,t}^{*} \qquad (9)$$

Panel A: Summary Statistics of the Distribution of the Estimated Coefficients

	$\alpha_0{}^*$	α_1	Γ	α_2	α_3	α_4	α_5	α_6	Mean R^2	Mean Obs.
mean(coef.)	-0.0026	0.5400	0.1725	0.0703	0.0432	0.0022	-0.0005	0.0040	0.5158	422
t(mean(coef.))	-5.4	21.6	13.2	13.9	9.4	6.8	-1.4	8.4		
median(t(coef.))	-1.2	11.2	3.0	3.5	2.0	1.1	-1.1	1.6		

The sample period includes the trailing four quarters of observations ending in quarter t for t = Q4:1997 through Q2:2015. Balance sheet items are measured at the end of the quarter. Income statement items are measured using the trailing four quarters of data. Details on variable definitions are provided in Section III and Appendix A. mean(coef.) is the time-series mean of the corresponding regression coefficient. t(mean(coef.)) is the t-statistic of the mean coefficient (the ratio of the time-series mean to its standard error). median(t(coef.)) is the time-series median of the regression t-statistic.

Panel B: Time-series Correlations between the Credit-loss-related Coefficients and Economy-wide State Variables

	α_1	γ	α_2	α_3	α_4	α_5	α_6
Pearson correlations							
%Δ in Real GDP	-0.31	-0.43	0.05	-0.15	-0.59	-0.20	0.20
p-value	0.009	0.0002	0.6567	0.2052	<.0001	0.0942	0.0985
Credit Spread (Baa-Aaa)	0.29	0.44	-0.05	0.07	0.47	0.22	-0.23
p-value	0.013	0.0001	0.6726	0.5521	<.0001	0.0699	0.0574
Spearman correlations							
%Δ in Real GDP	-0.09	-0.37	-0.05	-0.16	-0.52	-0.38	0.03
p-value	0.443	0.0017	0.6937	0.1757	<.0001	0.0009	0.8025
Credit Spread (Baa-Aaa)	0.03	0.39	0.09	0.17	0.60	0.48	-0.04
p-value	0.804	0.0007	0.4679	0.1556	<.0001	<.0001	0.7138

The table present coefficients and related p-values for the time-series correlations between the credit-loss-related coefficients (Table 2, Panel A) and economy-wide state variables, the percentage change in quarterly seasonally adjusted real gross domestic product relative to the same quarter a year ago (%Δ in Real GDP), and the difference between Moody's seasoned Baa and Aaa corporate bond yields (Credit Spread).

TABLE 3
Summary Statistics from Cross-sectional Regressions for Evaluating the Predictive Abilities of *ExpectedRCL* and NCOs

$$\frac{NCO_{i,t+1}}{AveLoans_{i,t+1}} = \beta_{0,t} + \beta_{1,t}\widehat{ExpectedRCL}_{i,t} + \beta_{2,t}\frac{NCO_{i,t}}{AveLoans_{i,t}} + \varepsilon_{i,t+1} \quad (11)$$

	β_0	β_1	β_2	Mean R^2	Mean Obs.
mean(coef.)	0.0007	0.9640		0.4722	428
t(mean(coef.))	3.3	26.4			
median(t(coef.))	1.2	19.6			
mean(coef.)	0.0019		0.7123	0.4115	428
t(mean(coef.))	9.3		21.3		
median(t(coef.))	5.9		17.3		
mean(coef.)	0.0008	0.9236	0.0131	0.4869	428
t(mean(coef.))	3.8	15.3	0.3		
median(t(coef.))	1.4	7.4	0.8		

The sample period includes the trailing four quarters of observations ending in quarter t for t = Q4:1997 through Q2:2014. Balance sheet items are measured at the end of the quarter. Income statement items are measured using the trailing four quarters of data. Details on variable definitions are provided in Section III and Appendix A. mean(coef.) is the time-series mean of the corresponding regression coefficient. t(mean(coef.)) is the t-statistic of the mean coefficient (the ratio of the time-series mean to its standard error). median(t(coef.)) is the time-series median of the regression t-statistic.

TABLE 4
Summary Statistics from Cross-sectional Regressions for Comparing the Predictive Abilities of *ExpectedRCL* and FVLoans

$$\frac{NCO_{i,t+1}}{AveLoans_{i,t+1}} = \beta_{0,t} + \beta_{1,t}\frac{FVLoans_{i,t}}{Loans_{i,t}} + \beta_{2,t}\widehat{ExpectedRCL}_{i,t} + \varepsilon_{i,t+1} \quad (12a)$$

$$\frac{NCO_{i,t+1}}{AveLoans_{i,t+1}} = \beta_{0,t} + \beta_{1,t}\frac{\widehat{FVLoans\_Macro}_{i,t}}{Loans_{i,t}} + \beta_{2,t}\frac{\widehat{FVLoans\_Other}_{i,t}}{Loans_{i,t}} + \beta_{3,t}\widehat{ExpectedRCL}_{i,t} + \varepsilon_{i,t+1} \quad (12b)$$

Panel A: Predictive Ability of *ExpectedRCL* and FVLoans Using Equation (12a)

	β_0	β_1	β_2	Mean R^2	Mean Obs.
mean(coef.)	0.0693	-0.0629		0.0860	241
t(mean(coef.))	6.5	-6.4			
median(t(coef.))	4.9	-4.4			
mean(coef.)	0.0011		0.8306	0.4477	241
t(mean(coef.))	2.3		17.1		
median(t(coef.))	0.1		13.7		
mean(coef.)	0.0133	-0.0122	0.8105	0.4532	241
t(mean(coef.))	4.6	-4.5	16.5		
median(t(coef.))	1.2	-1.1	12.6		

Panel B: Predictive Ability of *ExpectedRCL* and FVLoans Using Equation (12b)

	β_0	β_1	β_2	β_3	Mean R^2	Mean Obs.
mean(coef.)	0.0790	-0.0729	-0.0245		0.1014	241
t(mean(coef.))	6.3	-6.1	-2.2			
median(t(coef.))	4.3	-3.7	-0.79			
mean(coef.)	0.0011			0.8306	0.4477	241
t(mean(coef.))	2.3			17.1		
median(t(coef.))	0.1			13.7		
mean(coef.)	0.0165	-0.0155	-0.0087	0.8039	0.4576	241
t(mean(coef.))	3.9	-3.8	-1.1	17.1		
median(t(coef.))	1.1	-1.1	-0.2	12.4		

The sample includes 22 cross-sections (t): Q4:2005, Q4:2006, Q4:2007, Q4:2008, Q2:2009–Q3:2013. Balance sheet items are measured at the end of the quarter. Income statement items are measured using the trailing four quarters of data. Details on variable definitions are provided in Section III and Appendix A. mean(coef.) is the time-series mean of the corresponding regression coefficient. t(mean(coef.)) is the t-statistic of the mean coefficient (the ratio of the time-series mean to its standard error). median(t(coef.)) is the time-series median of the regression t-statistic.

TABLE 5
Summary Statistics from Cross-sectional Regressions for Comparing the Predictive Abilities of *ExpectedRCL*, the ALLL, and the PLLL

$$\frac{NCO_{i,t+1}}{AveLoans_{i,t+1}} = \beta_{0,t} + \beta_{1,t}\widehat{ExpectedRCL}_{i,t} + \beta_{2,t}\frac{ALLL_{i,t}}{Loans_{i,t}} + \beta_{3,t}\frac{PLLL_{i,t}}{AveLoans_{i,t}} + \varepsilon_{i,t+1} \quad (14)$$

	β_0	β_1	β_2	β_3	Mean R^2	Mean Obs.
mean(coef.)	0.0007	0.9640			0.4722	428
t(mean(coef.))	3.3	26.4				
median(t(coef.))	1.2	19.6				
mean(coef.)	-0.0013		0.4127		0.1723	428
t(mean(coef.))	-5.3		9.9			
median(t(coef.))	-0.9		7.9			
mean(coef.)	0.0012			0.6717	0.4381	428
t(mean(coef.))	8.9			27.0		
median(t(coef.))	4.0			18.1		
mean(coef.)	-0.0006	0.6046	0.1017	0.2707	0.5185	428
t(mean(coef.))	-5.9	20.9	7.4	14.0		
median(t(coef.))	-1.5	7.2	2.2	4.7		

The sample period includes the trailing four quarters of observations ending in quarter t for t = Q4:1997 through Q2:2014. Balance sheet items are measured at the end of the quarter. Income statement items are measured using the trailing four quarters of data. Details on variable definitions are provided in Section III and Appendix A. mean(coef.) is the time-series mean of the corresponding regression coefficient. t(mean(coef.)) is the t-statistic of the mean coefficient (the ratio of the time-series mean to its standard error). median(t(coef.)) is the time-series median of the regression t-statistic.

TABLE 6
Comparison of Absolute Prediction Errors in Forecasting Next Year's Ratio of Net Charge-offs to Average Loans

Panel A: All Bank Holding Companies

	Mean	Median	25th Ptl.	75th Ptl.	Std. Dev.	N	Mean Diff	Median Diff
$\widehat{ExpectedRCL}_t$	0.293%	0.145%	0.061%	0.332%	0.430%	27,312		
$ALLL_t/Loans_t$	0.383%	0.211%	0.096%	0.460%	0.496%	27,312	-0.091%***	-0.054%***
$PLLL_t/AveLoans_t$	0.306%	0.158%	0.068%	0.351%	0.437%	27,312	-0.013%***	-0.010%***
$NCO_t/AveLoans_t$	0.322%	0.148%	0.064%	0.369%	0.482%	27,312	-0.029%***	-0.007%***
$FVLoans_t/Loans_t$	0.518%	0.378%	0.182%	0.668%	0.525%	5,104	-0.164%***	-0.134%***

Panel B: Large Bank Holding Companies

	Mean	Median	25th Ptl.	75th Ptl.	Std. Dev.	N	Mean Diff	Median Diff
$\widehat{ExpectedRCL}_t$	0.304%	0.151%	0.061%	0.348%	0.439%	13,462		
$ALLL_t/Loans_t$	0.404%	0.207%	0.092%	0.490%	0.531%	13,462	-0.100%***	-0.053%***
$PLLL_t/AveLoans_t$	0.319%	0.154%	0.063%	0.367%	0.461%	13,462	-0.015%***	-0.007%***
$NCO_t/AveLoans_t$	0.338%	0.150%	0.064%	0.386%	0.502%	13,462	-0.034%***	-0.006%***
$FVLoans_t/Loans_t$	0.500%	0.368%	0.182%	0.646%	0.492%	2,994	-0.166%***	-0.133%***

Panel C: Small Bank Holding Companies

	Mean	Median	25th Ptl.	75th Ptl.	Std. Dev.	N	Mean Diff	Median Diff
$\widehat{ExpectedRCL}_t$	0.282%	0.141%	0.060%	0.318%	0.420%	13,850		
$ALLL_t/Loans_t$	0.363%	0.214%	0.101%	0.434%	0.459%	13,850	-0.081%***	-0.055%***
$PLLL_t/AveLoans_t$	0.293%	0.162%	0.073%	0.337%	0.412%	13,850	-0.012%***	-0.013%***
$NCO_t/AveLoans_t$	0.307%	0.146%	0.064%	0.353%	0.461%	13,850	-0.025%***	-0.008%***
$FVLoans_t/Loans_t$	0.545%	0.396%	0.181%	0.689%	0.567%	2,110	-0.162%***	-0.136%***

Panel D: High Commercial Real Estate Loans Concentration

	Mean	Median	25th Ptl.	75th Ptl.	Std. Dev.	N	Mean Diff	Median Diff
$\widehat{ExpectedRCL}_t$	0.299%	0.141%	0.058%	0.339%	0.443%	13,857		
$ALLL_t/Loans_t$	0.372%	0.211%	0.099%	0.432%	0.482%	13,857	-0.074%***	-0.052%***
$PLLL_t/AveLoans_t$	0.308%	0.157%	0.068%	0.347%	0.447%	13,857	-0.009%***	-0.008%***
$NCO_t/AveLoans_t$	0.321%	0.145%	0.063%	0.360%	0.494%	13,857	-0.022%***	-0.005%***
$FVLoans_t/Loans_t$	0.540%	0.381%	0.182%	0.682%	0.566%	2,705	-0.139%***	-0.117%***

Panel E: High Residential Real Estate Loans Concentration

	Mean	Median	25th Ptl.	75th Ptl.	Std. Dev.	N	Mean Diff	Median Diff
$\widehat{ExpectedRCL}_t$	0.281%	0.130%	0.052%	0.307%	0.444%	13,481		
$ALLL_t/Loans_t$	0.361%	0.189%	0.089%	0.408%	0.505%	13,481	-0.080%***	-0.049%***
$PLLL_t/AveLoans_t$	0.292%	0.147%	0.065%	0.313%	0.446%	13,481	-0.010%***	-0.011%***
$NCO_t/AveLoans_t$	0.308%	0.136%	0.060%	0.334%	0.489%	13,481	-0.027%***	-0.008%***
$FVLoans_t/Loans_t$	0.512%	0.360%	0.181%	0.651%	0.540%	2,626	-0.145%***	-0.120%***

Panel F: High Commercial and Industrial Loans Concentration

	Mean	Median	25th Ptl.	75th Ptl.	Std. Dev.	N	Mean Diff	Median Diff
$\widehat{ExpectedRCL}_t$	0.299%	0.156%	0.066%	0.338%	0.418%	14,142		
$ALLL_t/Loans_t$	0.399%	0.225%	0.099%	0.492%	0.496%	14,142	-0.100%***	-0.056%***
$PLLL_t/AveLoans_t$	0.320%	0.172%	0.072%	0.384%	0.432%	14,142	-0.022%***	-0.012%***
$NCO_t/AveLoans_t$	0.335%	0.161%	0.068%	0.397%	0.474%	14,142	-0.036%***	-0.008%***
$FVLoans_t/Loans_t$	0.536%	0.403%	0.193%	0.696%	0.522%	2,592	-0.199%***	-0.153%***

Panel G: Low Commercial and Industrial Loans Concentration

	Mean	Median	25th Ptl.	75th Ptl.	Std. Dev.	N	Mean Diff	Median Diff
$\widehat{ExpectedRCL}_t$	0.287%	0.135%	0.054%	0.325%	0.442%	13,170		
$ALLL_t/Loans_t$	0.367%	0.197%	0.094%	0.421%	0.496%	13,170	-0.081%***	-0.052%***
$PLLL_t/AveLoans_t$	0.291%	0.146%	0.065%	0.317%	0.442%	13,170	-0.004%*	-0.008%***
$NCO_t/AveLoans_t$	0.309%	0.136%	0.061%	0.342%	0.490%	13,170	-0.022%***	-0.006%***
$FVLoans_t/Loans_t$	0.500%	0.355%	0.175%	0.630%	0.527%	2,512	-0.129%***	-0.104%***

This table presents the distributional statistics of the absolute prediction errors of forecasting next year's ratio of net charge-offs to average loans using various credit risk metrics. The forecasts are the current year value of *ExpectedRCL* and the next year's predicted value of the ratio of net charge-offs to average loans estimated using ALLL, PLLL, NCOs, or FVLoans (Equation 15). Mean (Median) Diff is the mean (median) of the difference between absolute forecast errors based on *ExpectedRCL* and those based on other credit risk metrics. The significance of the mean (median) of the difference in absolute forecast errors is tested using t-tests (Wilcoxon signed-rank tests). Details on variable definitions are provided in Section III and Appendix A. *, **, and *** indicate p-values of less than 0.10, 0.05, and 0.01, respectively.

TABLE 7
Comparison of *ExpectedRCL* and Analyst Forecasts of the PLLL

	Mean	Median	25th Ptl.	75th Ptl.	Std. Dev.	N	Mean Diff	Median Diff
All bank holding companies								
$\widehat{ExpectedRCL}_t$	0.208%	0.137%	0.060%	0.269%	0.230%	1,344		
$Forecast_t$	0.423%	0.192%	0.075%	0.540%	0.570%	1,344	-0.215%***	-0.058%***
Large bank holding companies								
$\widehat{ExpectedRCL}_t$	0.179%	0.125%	0.057%	0.248%	0.178%	671		
$Forecast_t$	0.416%	0.186%	0.075%	0.585%	0.528%	671	-0.237%***	-0.071%***
Small bank holding companies								
$\widehat{ExpectedRCL}_t$	0.237%	0.151%	0.066%	0.296%	0.270%	673		
$Forecast_t$	0.429%	0.195%	0.074%	0.489%	0.609%	673	-0.192%***	-0.045%***
High residential real estate loans concentration								
$\widehat{ExpectedRCL}_t$	0.203%	0.132%	0.063%	0.267%	0.215%	671		
$Forecast_t$	0.408%	0.193%	0.071%	0.505%	0.566%	671	-0.205%***	-0.049%***
High commercial real estate loans concentration								
$\widehat{ExpectedRCL}_t$	0.233%	0.148%	0.067%	0.300%	0.263%	671		
$Forecast_t$	0.482%	0.234%	0.086%	0.629%	0.619%	671	-0.250%***	-0.082%***
High commercial and industrial loans concentration								
$\widehat{ExpectedRCL}_t$	0.203%	0.134%	0.056%	0.269%	0.228%	670		
$Forecast_t$	0.452%	0.203%	0.082%	0.641%	0.562%	670	-0.249%***	-0.089%***
Low commercial and industrial loans concentration								
$\widehat{ExpectedRCL}_t$	0.213%	0.138%	0.064%	0.269%	0.233%	674		
$Forecast_t$	0.393%	0.180%	0.067%	0.427%	0.577%	674	-0.180%***	-0.030%***

This table presents the distributional statistics of the absolute prediction errors of forecasting the provision for loan and lease losses using *ExpectedRCL* and the mean consensus analyst forecast of the provision for loan and lease losses (Forecast). The forecast errors are scaled by total assets and expressed as a percentage. Mean (Median) Diff is the mean (median) of the difference between absolute forecast errors based on *ExpectedRCL* and on the mean consensus analyst forecast. The significance of the mean (median) of the difference in absolute forecast errors is tested using t-tests (Wilcoxon signed-rank tests). Details on variable definitions are provided in Section III, Section VI, and Appendix A. *, **, and *** indicate p-values of less than 0.10, 0.05, and 0.01, respectively.

TABLE 8
Earnings Surprise Conditional on the Absolute Difference between *ExpectedRCL* and Forecasts of the PLLL and between *ExpectedRCL* and Forecasts of One-year-ahead NCOs

	Mean	Median	25th Ptl.	75th Ptl.	Std. Dev.	N	Mean Diff	Median Diff
Panel A: $Forecast_t$								
Difference below median	0.289%	0.150%	0.067%	0.305%	0.482%	662		
Difference above median	0.836%	0.409%	0.156%	1.024%	1.233%	668	-0.547%***	-0.259%***
Panel B: $NCO_t/AveLoans_t$								
Difference below median	0.442%	0.191%	0.080%	0.442%	0.778%	668		
Difference above median	0.686%	0.292%	0.119%	0.741%	1.130%	662	-0.245%***	-0.101%***
Panel C: $ALLL_t/Loans_t$								
Difference below median	0.510%	0.215%	0.084%	0.527%	0.962%	664		
Difference above median	0.616%	0.236%	0.110%	0.605%	0.989%	666	-0.106%**	-0.021%**
Panel D: $PLLL_t/AveLoans_t$								
Difference below median	0.492%	0.198%	0.087%	0.501%	0.952%	662		
Difference above median	0.634%	0.254%	0.105%	0.644%	0.996%	668	-0.142%***	-0.056%***

This table presents the distributional statistics of the absolute earnings surprise conditional on the absolute value of the difference between *ExpectedRCL* and (a) the mean consensus analyst forecast of the provision for loan and lease losses (Panel A) and (b) forecasts of next year's ratio of net charge-offs to average loans using the current year's ratio of net charge-offs to average loans (Panel B), allowance for loan and lease losses to total loans (Panel C), and provision for loan and lease losses to average loans (Panel D). The differences are partitioned based on whether they are greater than the median. Earnings surprise is calculated as the absolute value of the difference between realized earnings and the mean consensus analyst forecast of earnings, scaled by total assets. Mean (Median) Diff is the mean (median) of the difference between the earnings surprises across the median-based partitions. The significance of the difference in the mean (median) of the earnings surprise is tested using t-tests (Wilcoxon signed-rank tests). Details on variable definitions are provided in Section III, Section VI, and Appendix A. *, **, and *** indicate p-values of less than 0.10, 0.05, and 0.01, respectively.

TABLE 9
Using *ExpectedRCL* to Predict Bank Failure

$$\Pr(Fail = 1)_t = \beta_0 + \beta_1 \frac{Equity_t}{Assets_t} + \beta_2 \frac{Loans_t}{Assets_t} + \beta_3 \frac{RELoans_t}{Loans_t} + \beta_4 \frac{C\&ILoans_t}{Loans_t} + \beta_5 \frac{OREO_t}{Assets_t} + \beta_6 \frac{EarnNC_t}{Assets_t} + \beta_7 \frac{NPL_t}{Assets_t} + \beta_8 ROE_t + \beta_9 LIQ_t + \beta_{10} SIZE_t + \beta_{11} AGE_t + \beta_{12} \widehat{ExpectedRCL}_t + \beta_{13} \frac{ALLL_t}{Loans_t} + \beta_{14} \frac{PLLL_t}{AveLoans_t} + \beta_{15} \frac{NCO_t}{AveLoans_t} + \varepsilon_t \quad (16)$$

Panel A: Logit Regressions Predicting Bank Failure

	Column 1 **Logit model**	**Column 2** **Logit model**
$Equity_t/Assets_t$	-32.13***	-33.34***
$Loans_t/Assets_t$	7.00***	7.52***
$RELoans_t/Loans_t$	6.57***	6.60***
$C\&I\ Loans_t/Loans_t$	4.46	4.58
$OREO_t/Assets_t$	10.66	22.28
$EarnNC_t/Assets_t$	327.45***	312.41***
$NPL_t/Assets_t$	46.28***	14.04
ROE_t	-1.66***	-1.21*
LIQ_t	1.44	1.8
$SIZE_t$	0.07	0.05
AGE_t	-0.20	-0.25
$ALLL_t/Loans_t$	-70.04*	-34.2
$PLLL_t/AveLoans_t$	153.85***	94.87***
$NCO_t/AveLoans_t$	-168.01***	-229.59***
$\widehat{ExpectedRCL}_t$		139.51***
N	32,845	32,845
Failed banks	71	71
AIC	2345.48	2271.12

Panel B: Cox Proportion-hazard Models Predicting Bank Failure

	Column 1 Cox proportional-hazard model	Column 2 Cox proportional-hazard model
$Equity_t/Assets_t$	-30.95***	-30.48***
$Loans_t/Assets_t$	4.53***	5.23***
$RELoans_t/Loans_t$	3.57*	3.55*
$C\&I\ Loans_t/Loans_t$	2.66	3.23
$OREO_t/Assets_t$	-19.66	-13.31
$EarnNC_t/Assets_t$	250.47***	232.63***
$NPL_t/Assets_t$	59.54***	40.64***
ROE_t	-0.09	0.28
LIQ_t	0.34	1.64
$SIZE_t$	-0.11	-0.09
AGE_t	-1.08***	-0.86**
$ALLL_t/Loans_t$	-32.64	-21.45
$PLLL_t/AveLoans_t$	115.31***	84.53***
$NCO_t/AveLoans_t$	-99.41***	-147.54***
$\widehat{Expected}RCL_t$		98.95***
N	32,845	32,845
Failed banks	71	71
AIC	507.44	495.04

This table provides the results of logit regressions (Panel A) and Cox proportional-hazard models (Panel B) predicting bank failure using various bank-specific variables including *ExpectedRCL* and other credit risk metrics. In the logit model, standard errors are clustered by bank. Details on variable definitions are provided in Section III, Section VI, and Appendix A. *, **, and *** indicate p-values of less than 0.10, 0.05, and 0.01, respectively.

TABLE 10
Summary Statistics from Cross-sectional Regressions Comparing the Predictive Abilities of *ExpectedRCL*, NCO, ALLL, PLLL, and FVLoans for Long-horizon Net Charge-offs

$$\frac{\sum_{i=1}^{3} NCO_{i,t+i}}{\sum_{i=1}^{3} AveLoans_{i,t+i_t}} = \beta_{0,t} + \beta_{1,t}\widehat{ExpectedRCL}_{i,t} + \beta_{2,t}\frac{ALLL_{i,t}}{Loans_{i,t}} + \beta_{3,t}\frac{PLLL_{i,t}}{AveLoans_{i,t}} + \beta_{4,t}\frac{NCO_{i,t}}{AveLoans_{i,t}} + \beta_{5,t}\frac{FVLoans_{i,t}}{Loans_{i,t}} + \varepsilon_{i,t+1} \quad (17)$$

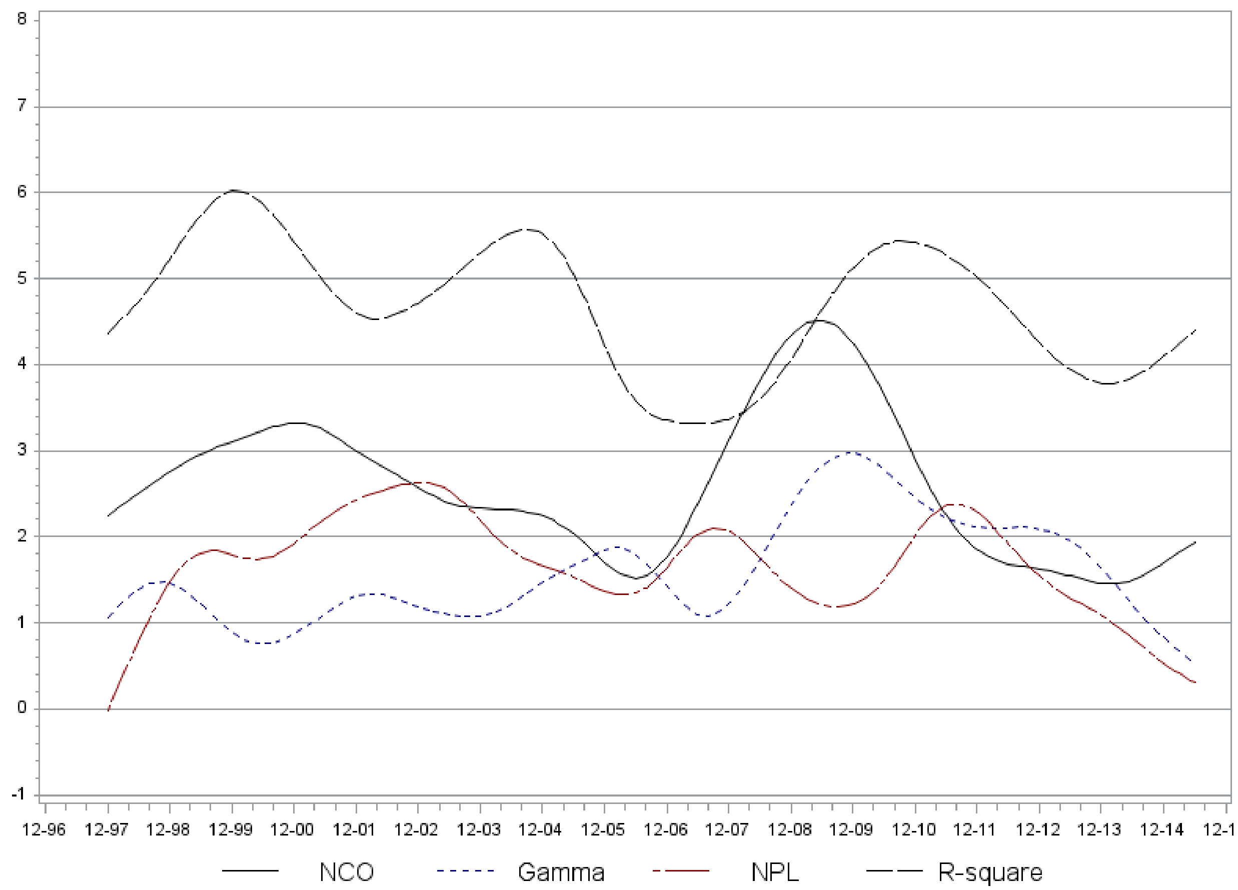
	β0	β1	β2	β3	β4	B5	Mean R^2	Mean N
mean(coef.)	0.0022	0.8707					0.3832	362
t(mean(coef.))	6.9	14.9						
median(t(coef.))	4.7	16.0						
mean(coef.)	0.0007		0.3435				0.1217	362
t(mean(coef.))	3.4		9.9					
median(t(coef.))	1.0		5.8					
mean(coef.)	0.0025			0.6201			0.3460	362
t(mean(coef.))	8.3			15.1				
median(t(coef.))	5.6			14.3				
mean(coef.)	0.0032				0.6325		0.3207	362
t(mean(coef.))	9.0				15.9			
median(t(coef.))	8.6				13.7			
mean(coef.)	0.0523					-0.0452	0.0773	207
t(mean(coef.))	6.61					-6.019		
median(t(coef.))	4.43					-3.87		
mean(coef.)	0.0012	0.8014	0.0633	0.3593	-0.3492		0.4289	362
t(mean(coef.))	4.8	10.7	5.2	6.6	-5.0			
median(t(coef.))	1.1	5.2	1.0	3.2	-1.6			
mean(coef.)	0.0054	0.5546	0.0959	0.2146	-0.2646	-0.0034	0.4360	207
t(mean(coef.))	1.92	5.99	2.19	2.92	-2.89	-1.24		
median(t(coef.))	0.47	5.00	0.55	2.41	-1.51	-0.26		

The sample period includes the trailing four quarters of observations ending in quarter t for t = Q4:1997 through Q2:2012. The sample period for models with FVLoans includes Q4:2005, Q4:2006, Q4:2007, Q4:2008, and Q2:2009–Q2:2012. Balance sheet items are measured at the end of the quarter. Income statement items are measured using the trailing four quarters of data. Details on variable definitions are provided in Section III and Appendix A. mean(coef.) is the time-series mean of the corresponding regression coefficient. t(mean(coef.)) is the t-statistic of the mean coefficient (the ratio of the time-series mean to its standard error). median(t(coef.)) is the time-series median of the regression t-statistic.

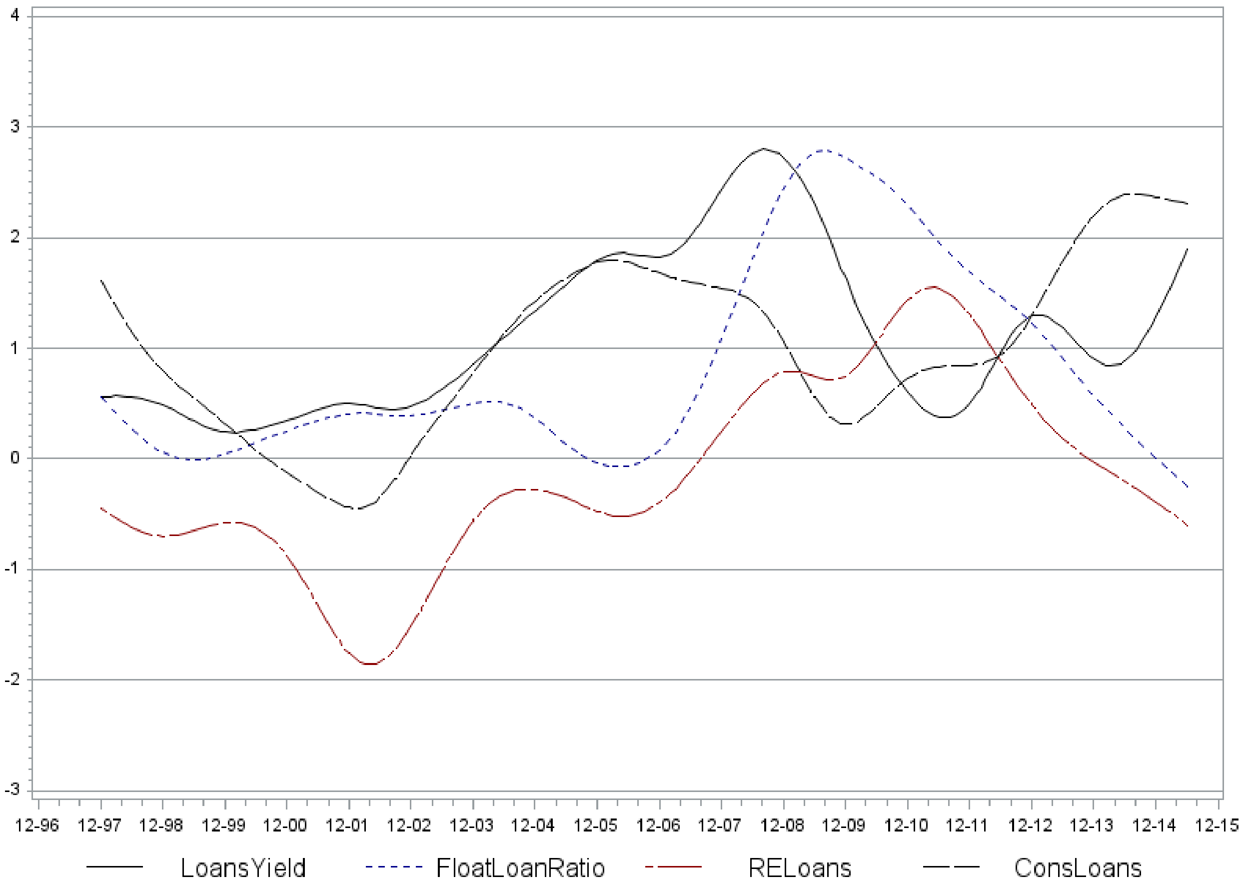
FIGURE 1
Standardized Coefficients from Cross-sectional Regressions for Estimating *ExpectedRCL*

$$\frac{NCO_{i,t}}{AveLoans_{i,t}} = \propto_{0,t}{}^{*} + \propto_{1,t} \frac{NCO_{i,t-1}}{AveLoans_{i,t-1}} + \propto_{1,t} \gamma_t \frac{\Delta NPL^{unexp}_{i,t-1}}{AveLoans_{i,t-1}} + \propto_{2,t} \frac{NPL_{i,t-1}}{Loans_{i,t-1}} + \propto_{3,t} LoansYield_{i,t-1} + \propto_{4,t} FloatLoanRatio_{i,t-1} + \propto_{5,t} \frac{RELoans_{i,t-1}}{Loans_{i,t-1}} + \propto_{6,t} \frac{ConsLoans_{i,t-1}}{Loans_{i,t-1}} + \varepsilon^{*}_{i,t} \quad (9)$$

Panel A: Coefficients for Credit-loss Variables and R-squared



Panel B: Other Coefficients



The figure presents standardized coefficients for the reported variables and R-squared from cross-sectional regressions of Equation (9) over time. To ease interpretation, the coefficients (and R-squared) are standardized by dividing them by their time-series standard deviation. The figure plots a smoothed version of the coefficients. The sample period includes the trailing four quarters of observations ending in quarter t for t = Q4:1997 through Q2:2015. Balance sheet items are measured at the end of the quarter. Income statement items are measured using the trailing four quarters of data. Details on variable definitions are provided in Section III and Appendix A.