Sentiment in Bank Examination Reports and Bank Outcomes

Maureen Cowhey[*], Seung Jung Lee[†], Thomas Popeck Spiller[‡], Cindy M. Vojtech[§¶]

October 11, 2022

Abstract

We investigate whether the bank examination process provides useful insight into bank future outcomes. We do this by conducting textual analysis on about 5,500 small to medium-sized commercial bank examination reports from 2004 to 2016. These confidential examination reports provide textual context to the components of supervisory ratings: capital adequacy, asset quality, management, earnings, and liquidity. Each component is given a categorical rating, and each bank is assigned an overall composite rating, which are used to determine the safety and soundness of banks. We find that, controlling for a variety of factors, including the ratings themselves, the sentiment supervisors express in describing most of the components predict relevant future bank outcomes. The sentiment conveyed in the asset quality, management, and earnings sections provides significant information in predicting future outcomes for problem loans, supervisory actions, and profitability, respectively, for all banks. Sentiment conveyed in the capital adequacy section appears to be predictive of future capital ratios for weak banks. These relationships suggest that bank supervisors play a meaningful role in the surveillance of the banking system.

JEL classification: G21; G28.

Keywords: Bank examination reports; CAMELS; private supervisory information; Natural Language Processing.

---

[*]E-mail: maureencowhey@outlook.com

[†]E-mail: seung.j.lee@frb.gov

[‡]E-mail: thomas.w.spiller@frb.gov

[§]E-mail: cindy.m.vojtech@frb.gov

[¶]Maureen Cowhey is at the University of California, Los Angeles; all other authors (Seung Jung Lee, Thomas Popeck Spiller, and Cindy M. Vojtech) are at the Board of Governors of the Federal Reserve System. We thank Matthew Deininger, Nathan Mislang, Ariel Polani, Mariya Pominova, Sarah Ridout, Ashley Sexton, Krista Stapleford, and Nicholas Stewart for help cleaning the data and for research assistance. We thank participants at the Financial Stability Workshop and the International Banking and Finance Lunch Brownbag for their comments, and we are grateful for Tamas Vadasz for helpful suggestions as a discussant at the 2022 IBEFA conference. Disclaimer: The material here does not represent the views of the Board of Governors of the Federal Reserve System or its staff.

1 Introduction

This paper investigates the effectiveness of supervision in maintaining the safety and soundness of the banking system in the United States. Supervisors play a vital role in the surveillance of banks by actively engaging in bank examinations and off-site monitoring. Many studies on the effectiveness of bank supervision have relied on the categorical CAMELS ratings of banks' safety and soundness. We take a close look at additional information—the examination reports based on on-site exams of small and medium-sized banks—to gauge the usefulness of more detailed information in this surveillance process. In particular, we study how sentiment in these reports conveys important information on future bank outcomes.

Banks are important intermediaries in the U.S. financial system. While offering a plethora of financial services, their main function is to receive and manage deposits in order to originate loans and invest in securities. Partly due to the maturity mismatch between assets (which are more long-term) and liabilities (which are more short-term), banks are susceptible to runs. Bank runs, in turn, could render the entire financial system unstable. Deposit insurance helps to prevent some of these unwarranted runs but creates limited liability for its stakeholders which may motivate excess risk-taking. Risk taking also has financial stability implications, if occurring on a large scale. Therefore, banks are subject to various forms of capital requirements that ensure banks' ability to absorb nontrivial shocks to their earnings and balance sheets.

In this context, bank supervision, along with regulation, plays a key role in maintaining the banking system's safety and soundness by establishing a good understanding of banks' capital adequacy, asset quality, management effectiveness, earnings prospects, liquidity positions, and sensitivity to market risk. Supervisors convey their assessments on these measures through CAMELS ratings. These ratings summarize both public data and private supervisory information gathered during on-site bank exams and have been found to contain information useful to the supervisory monitoring of commercial banks (Berger and Davies (1998) and Berger, Davies and Flannery (2000)).

However, past studies have relied on these categorical ratings (of integers from 1 to 5) in showing the effectiveness of supervision and not from the actual content of the bank exams. These reports contain the insights and understanding gained from the bank examination process. To achieve a more thorough and complete understanding of the usefulness of the examination process, it is important to look at whether the reports also convey extra information beyond what is conveyed in the discrete ratings.

In this paper, we look at about 5,500 bank examination reports from 2004 to 2016 and calculate sentiment scores based on the text in these documents. We calculate the sentiment on the language associated with five of the components of CAMELS: capital adequacy, asset quality, management, earnings, and liquidity.[1] We use this information to see if it has any additional predictive power for determining various bank outcomes related to each component. We also test the informational content of sentiment given a particular setting. For example, we compare results by weak CAMELS score or strong score, by splitting the sample into the financial crisis period (up to 2011) or post-crisis, and by testing at different time horizons.

We find that the sentiment conveyed in different components of the CAMELS ratings have varying impacts in different circumstances. In particular, we find that controlling for a variety of factors, the sentiment supervisors express in describing many of the components predict relevant future bank outcomes. More specifically, the sentiment conveyed in the asset quality, management, and earnings sections provides significant information in predicting future outcomes for problem loans, supervisory actions, and profitability, respectively. This predictive relationship in turn is driven by banks with better ratings when it comes to management, and banks with worse ratings when it comes to asset quality and earnings. Moreover, we find that the sentiment conveyed in the capital adequacy section of the exams are predictive of future capital ratios for weak banks. These relationships suggest that bank supervisors play a meaningful role in the surveillance of the banking system, focusing on different aspects of safety and soundness in different circumstances.

The rest of the paper is as follows. Section 2 provides a literature review, and section 3 provides a more detailed description of the bank examination process. Section 4 briefly describes the data, and section 5 investigates which sentiment score to use for our analysis. Then we provide our main econometric specification and results, followed by a conclusion.

2 Literature

Our paper mainly contributes to two strands of literature. The first strand is related to the information created though the bank examination process, mainly through the determination of supervisory ratings and related enforcement actions. Measuring the effectiveness of these activities have generally relied on ratings and enforcement actions. Our paper also contributes to the more recent and growing literature on extracting sentiment information from financial text and testing if such sentiment matters for predicting macroeconomic or financial outcomes.

Supervisory safety and soundness bank ratings and related supervisory actions have been shown to have useful private information. For example, most recently Gaul and Jones (2021) find that

[1]Sensitivity to market risk was added to the framework in 1995. The text associated with this score is not widely available in the sample of banks we have, so we skip the analysis on this particular category. Smaller banks tend to have a limited trading book, so market risk would be of less relevance to these banks.

CAMELS ratings (and the Management component rating) have significant predictive power for future bank performance and risk measures relevant to bank regulators and supervisors. DeYoung et al. (2001) show that on-site commercial bank examinations produce value-relevant information about the future safety and soundness of banks reflected in bond prices of parent holding companies. Berger, Davies and Flannery (2000) note that supervisory assessments immediately following a bank examination generally contribute substantially to forecasting future problem loans and bank earnings, often exceeding the contribution of market assessments. In addition, Berger and Davies (1998) show that CAMELS rating downgrades have a significant relationship with abnormal returns. Similarly, Jordan, Peek and Rosengren (2000) show that the announcement of formal supervisory actions have stock market reactions; only banks in the same region as the announcing bank, with similar exposures, are found to be affected. Finally, at a more aggregate level, going beyond just looking at publicly listed large commercial banks or bank holding companies, Peek, Rosengren and Tootell (1999) show that the percentage of commercial bank assets associated with the worst CAMELS ratings helps provide more accurate forecasts of macroeconomic variables such as the unemployment rate and inflation than can be predicted by Federal Reserve Board staff. Indeed, supervisory ratings are also used to determine FDIC deposit insurance premiums and examination frequency, given their high informational content. Our paper tries to add to this literature by providing more granular evidence of private information creation during the bank examination process. This involves extracting additional information from the bank examination reports through sentiment scoring.

Most research on bank supervision has focused on bank ratings and supervisory actions. A recent exception is Hirtle, Kovner and Plosser (2020) which uses data on supervisory hours. The authors find that banks which receive more supervisory attention hold less risky loan portfolios and are less sensitive to industry downturns. Yet, these same banks do not have lower growth or profitability. These results suggest that supervisors help mitigate risk in the financial system while not undermining competitiveness.

Up to now, sentiment analysis in the economics and finance literature has generally been used on mainly three different types of publicly available text: economic news, central bank communications, and corporate financial filings or earnings calls. Recent papers have analyzed new ways to combine commonly used lexicons with machine learning techniques to construct sentiment scores that accurately extract signal from economic news text to predict future changes in macroeconomic and financial cycle indicators. (See Nyman, Kapadia and Tuckett (2021), Shapiro, Sudhof and Wilson (2020), and Kalamara et al. (2020).) In addition, other research combines these same techniques and utilizes them to determine how central bank communications (in the form of internal and external reports, FOMC meetings, and/or internal communications) accurately predict financial crises and future policy decisions or drive changes in certain macroeconomic indicators, such as inflation. (See Correa et al. (2021), Shapiro and Wilson (2019), and Hubert and Labondance (2017).) Most relevant to our work are the papers that look at earnings call transcripts and corporate financial filings. For example, several papers such as Jiang et al. (2019) and Price et al. (2012) show that the tone of the earnings calls lead to significant changes in stock market prices. More relevant to the banking industry, a few papers build off of Loughran and McDonald (2011) by using their finance lexicon to look at how sentiment expressed by corporate managers in official filings can predict future financial distress at banks. (See Gandhi, Loughran and McDonald (2019), Nopp and Hanbury (2015), and Gupta et al. (2022).)

Our analysis seeks to expand the literature on sentiment analysis by looking at the relationship between the sentiment expressed in bank examination reports used in supervisory assessments and a bank's CAMELS ratings, as well as a host of other quantitative factors that are traditionally indicative of a bank's financial soundness. While Goldsmith-Pinkham, Hirtle and Lucca (2016) use computational linguistic methods to categorize Matters Requiring Attention—one text product of bank supervision reports—into a number of topics, our study differs in that we calculate a sentiment score covering all of the text, and we do so for the full-scope report that accompanies the CAMELS ratings.

3 Bank Examination Process

Every commercial bank undergoes a comprehensive bank exam about once a year. This process is led by one bank regulator, and sometimes in partnership with another regulator. In the case of State-Member Banks (SMB), examinations are performed alternating leadership between the Federal Reserve Bank responsible for the SMB and the state-level financial regulator. The goal of these "full scope" exams is to assess the safety and soundness of a commercial bank by reviewing any problems that were identified last round, scoring the bank on six categories of safety and soundness, and generating an overall composite score. This scoring system is called CAMELS and is an acronym of the subscores:

- Capital Adequacy: Representing the ability of the bank to absorb losses
- Asset Quality: Representing the known and likelihood of losses the bank might face
- Management: Representing the quality of the management team, compliance function, audit function, and business strategy
- Earnings: Representing the ability of the bank to provide returns on their activities
- Liquidity: Representing the ability of the bank to absorb short term funding difficulties

- Sensitivity to Market Risk: Representing the bank's exposure to markets such as interest rate changes and marketable securities

Each subscore and the overall composite is rated from 1 (strongest) to 5 (weakest). And depending on the composite rating, the period between comprehensive exams will change. Banks with ratings of 3, 4, or 5 are considered "weak" banks or banks with weak ratings and banks with ratings of 1 or 2 are considered "strong" banks or banks with strong ratings. Weak banks are examined every 6 months, and strong banks may have their examinations up to once every 18 months.[2]

To finish the exam, the examination team writes a Report of Examination, which is vetted through the leadership of the regulator, and the team presents their findings to the executive team of the bank. Because of this vetting and presentation, the language of the examination is important. The examiner must be able to justify their rating both with financial ratios and other information, as explored in Bassett, Lee and Spiller (2015), and in the text of the examination.

The CAMELS ratings themselves are private supervisory information and weaker ratings mean that banks are subject to, not only more frequent examinations, but also restrictions on certain activities such as mergers and acquisitions, dividend payouts, and new activities. However, these ratings have discrete integer values that range from 1 to 5. Because the text in examination reports are used to justify these ratings, the text should be highly correlated with the ratings. Any additional informational content in these examination reports that help predict future bank outcomes would signify that supervisors produce more granular insights into the safety and soundness of banks than can be gauged by the CAMELS ratings alone.

To provide an overview of the sample data, figure 1 shows the number of exams for each quarter in the sample for which we have examination reports. We restrict banks to those with less than \$10 billion in assets for our analysis. Since the late 2000s, we have bank examination data from about 100 to 150 bank examination per quarter, which makes up from 75 to 80 percent of all possible exams. Earlier in our sample, this number is about 50 to 100 per quarter due to data limitations, which composes far less (about 20 to 50 percent of possible exams). Figure 2 shows the distributions for the component and composite CAMELS ratings. Most banks are rated strong at a given point in time, but over the financial crisis many banks received weak ratings for their component scores and their composite scores, particularly for earnings and asset quality.

[2]Prior to The Federal Deposit Insurance Corporation Improvement Act of 1991 (FDICIA) effective in December 1992, the three federal regulators and many state banking departments examined banks on frequencies that varied by banks condition or past ratings. FDICIA established a uniform criteria based on size and risk profile, which has, since then, changed over time. The threshold for the risk profile last changed as part of The Riegle Community Development and Regulatory Improvement Act of 1994, and the threshold for size last changed in 2007 as part of the Interim Rules to Implement the Examination Amendments of 2006, which require annual examinations of banks with assets greater than \$500 million. See Rezende (2014) for more details.

4 Extracting Sentiment from Bank Exams

4.1 Sentiment Score Design

This section evaluates three lexicons and three methods to characterize the sentiment of the text within the bank exam reports. A lexicon refers to a pre-defined list of words assigned to positive or negative values, classifying each word with positive or negative sentiment, respectively. The three methods use different techniques to weight those positive and negative values.

Proper selection of the lexicon is important. As Loughran and McDonald (2011) highlight in their construction of a new finance-based lexicon, words often have different meanings in different domains. For example, as insinuated in the title of their seminal piece on the issue, while "liability" is generally considered a negative word in English vernacular, in finance or economics, the word is neutral. As a result, we explore three different lexicons to construct scores for the sentiment expressed by bank examiners. The Loughran and McDonald (LM) lexicon was created specifically for finance and economics-related text. The Financial Stability (FS) lexicon (Correa, Garud, Londono and Mislang, 2021) was created to analyze the sentiment of central bank financial stability reports. Finally, Hu and Liu's (Hu and Liu, 2004) opinion lexicon (QDAP)was constructed with broader sentiment analysis in mind and are not tailored to specific text in finance/economics.[3]

We apply these dictionaries using three methods. The first sentiment score approach is oftentimes called referred to as the bag-of-words approach, as it neglects the role that grammar, syntax, and other context-specific traits play in the overall sentiment of text and merely bases the analysis on individual words. The score is based on the following index calculation:

$$\text{sentiment index}_{i,t} = \frac{\#PositiveWords - \#NegativeWords}{\#NegativeWords + \#PositiveWords},$$

where i represents either the report as a whole or the subsections of the report that correspond to the different components of the CAMELS rating.

However, merely utilizing a raw word count is likely insufficient; a particular word's frequency is directly related to the length of the document, not its importance. Moreover, a word that occurs infrequently likely carries more weight to the average reader than a word that is commonplace in the document, and the raw word count does not account for this. To address this, we apply a sentiment score method that adjusts the sentiment scores for term frequency and inverse document frequency, a weighting mechanism commonly referred to as tf-idf. Terms that are used frequently

[3]Other lexicons include the Harvard-IV-4 psychosocial dictionary that is also broad-based, and Elaine Henry's (Henry, 2008, HE) lexicon was created to develop sentiment measures for earnings press releases.

are down-weighted using these formulas

$$\begin{aligned}
\text{term frequency: } tf_{t,d} &= \frac{\text{number of term}}{\text{total words}}, \\
\text{inverse term frequency: } idf_{t,D} &= log\Big(\frac{N}{|d \in D : t \in d|}\Big), \\
tfidf_{t,d,D} &= tf_{t,d} * idf_{t,D},
\end{aligned}$$

where N is the number of documents in the corpus and $|d \in D : t \in d|$ is the number of documents where the term t appears (i.e., $tf(t,d) \neq 0$).

Finally, in order to overcome the bag-of-words method's failure to incorporate context into its calculations, we consider a third method that also takes valence shifters into account. Valence shifters are words that either change the polarity of the word (i.e., negators like "not" and "never") or amplify the word (i.e., words like "slightly" or "extremely"). We use an R package called -sentimentR-, which is designed to consider four words before and two words after a dictionary-weighted word to search for valence shifters and to weight words differently based on their association with various amplifiers.

4.2 Evaluating Sentiment Scores

Because there are numerous ways to capture sentiment, the first step is to select a few methods that seem best suited to capture the information content of a bank exam. To assess the performance of the sentiment scores, we compare how well sentiment measures explain CAMELS scores. In particular, we regress CAMELS scores on sentiment. We do this for the composite score and sentiment scores built using all of the text, and also for scores of CAMELS components and sentiment scores based on only that section of the exam text. The exact specification is

$$\text{CAMELS score}_s = \alpha + \beta \text{Sentiment score}_{s,m,l} + \epsilon_s,$$

where s is the section of the exam, m is the sentiment score method, and l is the sentiment score lexicon, with constant term α, coefficient β, and an error term ϵ.

Table 1 summarizes the results by reporting the R-squared and adjusted R-squared for each specification. The three methods are listed in the first column: polar, weighted polar (tf-idf), and valence. The second column is the bank exam section of text used, and the three lexicons are listed across the top: FS, LM, and QDAP. The best performer based on R-squared or adjusted R-squared is listed in the last two columns.

Note that, except for the liquidity portion of the exams, the variation in CAMELS ratings can be relatively well explained by any of the sentiment measures calculated by the different methods used. The R-squareds for using the sentiment measures from the LM dictionary, in particular, explain about 10 to 35 percent of the variation in the CAMELS ratings. Indeed, the LM lexicon is the best performer overall. In addition, comparing the level of the R-squared or R-squared score, the regular polar generally performs better than the weighted polar (tf-idf). The valence method generally performs better than the regular polar. We retain the two best methodologies to use for the main analysis: LM regular polar and LM valence, in part because this implies that these measures are best reflective of sentiment embedded in these examination reports, but also because we want to be conservative in trying to find extra informational content (beyond the CAMELS ratings themselves) of sentiment in explaining future bank outcomes in sections 7 and 8.

Indeed, a simple chatterplot example in figure 3 illustrates that the LM dictionary has nice differentiating properties when it comes to the language used in bank exam reports. Positive words (blue) in the LM dictionary are used more often in the Earnings section of bank exam reports associated with stronger ratings (score of 1 or 2), while negative words (red) in the LM dictionary are more frequently used in bank exam reports with weaker ratings (score of 3, 4, or 5). For other sections or for the entire bank examination document, such differentiation is evident as well (see Appendix).

To provide a better idea of how the LM polar and LM valence-based sentiment scores are associated with CAMELS ratings, figure 4 shows that in general these two methods have differentiating properties when it comes to stronger vs. weaker banks as can be seen by the distinct humps in the kernel densities of sentiment scores that vary with how strong or weak the composite CAMELS scores are. We show that a more granular distinction of stronger vs. weaker banks also relates to helping predict bank outcomes. Note that the polar sentiment score ranges from negative one to positive one, but usually reside in negative territory, meaning that there are more negative words than positive words used in the bank examination reports. The weighting in the valence method makes the range of that score a bit narrower.

5 Data

Our examination data primarily consist of SMBs because SMBs are directly regulated by the Federal Reserve. To make the sample more internally comparable, we drop the largest banks, specifically commercial banks that are a part of a bank holding company that is required to be stress tested frequently. Practically speaking, we eliminated any banks with more than $10 billion in assets. These larger banks are subject to a much more rigorous set of regulations and more oversight.

In addition to the text of the bank exams, we also have information on financial performance and supervisory activity. More specifically, we have counts of how many times regulators cited matters requiring attention (MRA) or matters requiring immediate attention (MRIA) in a bank exam. The remaining bank performance data come from Consolidated Reports of Condition and Income, also referred to as the Call Reports, which are quarterly financial statements that commercial banks are required to report (FFIEC 031 and 041).

Within the bank exam report, the data of the financial statements used is indicated. We use that date for merging the exam and Call Report data. We test financial ratios that map to the CAMELS categories. For example, we look the Tier 1 ratio and the common equity Tier 1 (CET1) ratio for capital adequacy outcomes; loan loss provisions to loans, net charge-offs to assets, and delinquent loans to loans for asset quality outcomes; return on assets (ROA) and pre-provision net revenue (PPNR) to assets for earnings outcomes, and securities or cash and securities to assets for liquidity outcomes. For ratios with flow variables in the numerator, we use the flow data over the past four quarters. The denominator is averaged over the past four quarters in order to get a weighted ratio. We use MRAs and MRIAs for the composite and management outcomes. The final sample used for testing has about 5,500 observations between 2004:Q1 and 2016:Q2, which accounts for about 60 percent of all possible examinations. After 2008, we have data that accounts for about 75 to 80 percent of all possible examinations.

Table 2 summarizes the data. The top panel reports the sentiment scores that are discussed in the previous section. Recall that the sentiment score construction creates a value ranging between -1 and 1. In order to simplify interpretation of regression results, we normalize the scores to range between 0 and 1. This has no bearing on our results from our main econometric specification, but is important for interpretation when we interact the sentiment scores with other variables.

The bottom panel reports the control variables, some of which also serve as outcome variables. To control for outliers, all the various financial ratios are winsorized at 1 and 99 percent. However, to maintain confidentiality, we show the 5th and 95th percentiles of all the variables in our sample in table 2. Bank assets range between about \$32 million and \$3 billion when it comes to the 5th and 95th percentiles.

6 Econometric Specification

This section will explain the testing design. The goal is to pull out the information embedded in the textual content of bank exams. In particular, model specifications relate the soft information in exams to future bank outcomes above and beyond the CAMELS ratings the text justifies. Section 4.2 evaluated how well the various sentiment methods captured the CAMELS ratings. Now that the two sentiment score methods have been established, we will test to see what additional information, beyond the CAMELS score, is contained in the text.

For each CAMELS score category, we identify performance measures related to that category: capital, asset quality, management, earnings, liquidity, and the composite score. We then test to see if sentiment from an exam is related to a performance measure one year later after controlling for other observables, including the ratings themselves. Any statistically significant results that link sentiment to bank outcomes will be that much more profound, as we use the sentiment methods that are the most highly correlated with the CAMELS scores.

The first set of regressions uses the full sample between 2004:Q1 and 2016:Q2. Additional tests will subset the data in various ways. The baseline specifications follow this structure:

$$\begin{aligned} \text{outcome}_{i,t} &= \rho \,\text{outcome}_{i,t-1} + \beta \,\text{sentiment}_{i,c,t-1} + \gamma \, log(\text{assets}_{i,t-1}) \\ &\quad + \Sigma_{n=1}^{4} \, \psi_n \, \text{CAMELS dummy}_{i,c,n,t-1} + \phi_t + \epsilon_{i,t}, \end{aligned}$$

for bank i, in period t, for bank exam component c, and where ϕ_t is a time fixed effect. Table 2 lists the outcome variables for each of the exam sections. For data that comes from the Call Report (all variables except MRA/MRIA), the regression is testing outcomes one year after the exam. Because MRAs/MRIAs get completely refreshed at a "full scope" exam, these regressions are run using adjacent exams. Exams are generally 6 to 18 months apart.

The lagged value of the outcome variable (value at the time of the exam with the sentiment score) is included because many of these performance metrics are persistent, and again, in order to assess the information content of the exam text, we are trying to strip out simple observable information. This dynamic specification prevents us from using bank fixed effects that would bias coefficient estimates, but our results are qualitatively and quantitatively simillar if we included bank fixed effects (not shown).[4] The coefficient of interest is β, the loading on the sentiment score. Notice that the sentiment score is subscripted by c, the exam component. The sentiment score can be based on all of the text in the case of testing with the composite CAMELS score, or the sentiment score could be based on text from one component such as asset quality that is then tested with the asset quality CAMELS score. By including the set of relevant CAMELS score dummies, we are controlling for the general performance of the bank. The specification design is testing whether the information content of the exam text is associated with future bank outcomes.

[4]The main difference is that in the bank fixed effects specification, the sentiment in the capital adequacy sections of the exam has a statistically significant effect in predicting future capital ratios. Without bank fixed effects, this result only holds for the weak banks in the sample, which we show later.

7 Main Results

This section provides a description of the results of our baseline regressions by different categorical ratings, including the composite.

Composite

Table 3 shows the results of the regressions using a sentiment score based on all of the exam text. One outcome variable is based on the number of MRA and MRIA. The second outcome variable takes a value of one if there are any MRAs/MRIAs and a value of zero otherwise. More of these “matters” are a sign that the bank is not operating properly and has supervisory concerns requiring attention. The coefficient on the lag sentiment score is negative as expected and is highly significant across all specifications. Higher or more positive sentiment in the exam text is associated with a lower number of MRA/MRIAs at the next comprehensive exam. To put the magnitude of the coefficient in perspective, a one-standard deviation change in either the polar or valence LM sentiment (0.13 and 0.06, respectively) leads to about a decrease of 1 MRA or MRIA or a decrease in the likelihood of an MRA or MRIA by about 10 percentage points. Given that there are about three MRA/MRIAs per exam on average, these are significant results. In fact, in terms of standard deviations, this is equivalent to a decrease about 20 percent of a standard deviation in the sum of MRA/MRIAs or the likelihood of MRA/MRIAs being issued. Lagged MRAs or MRIAs and CAMELS composite ratings are also highly relevant to the future number and likelihood of MRAs or MRIAs. The CAMELS dummy coefficients generally increase monotonically as the CAMELS ratings become worse. The lowest quality score (a five) is associated with more MRA/MRIAs, but the coefficient is sometimes slightly less than the coefficient on the four dummy.

Capital Adequacy

Table 4 shows the results of using the text in the capital adequacy section. The coefficient of interest, the lagged sentiment score, is positive but not statistically significant across all specifications. These relatively small effects compared to the standard deviation of capital ratios themselves (both over 10 percentage points) are consistent with a bank's capital adequacy status being well-informed simply by using regulatory ratios. This is also indicative from the high R-squareds of about 0.8. Interestingly, only when the CAMELS capital-specific rating is a five, does the rating have a statistically significant and a significant quantitative effect on future capital ratios. We investigate whether for the weakest banks, sentiment conveyed in the capital adequacy section of bank examination reports matters in section 8.

Asset Quality

Table 5 shows the results of testing measures of asset quality. Three ratios are tested: loan loss provisions, four quarters of net charge-offs (NCOs), and delinquencies. All measures are defined relative to all loans. Note that in all cases a higher ratio is associated with worse asset quality.

The coefficient of interest on lag sentiment is negative and statistically significant across all specifications, indicating that better sentiment is associated with improved asset quality one year later. A one-standard deviation improvement in either the polar or valence LM sentiment in the asset quality section of bank exams (0.16 and 0.08, respectively) leads to about 4 to 5 basis point decrease in loan loss provision rates, about 1 to 2 basis point decrease in net charge-off rates, and more than about 13 to 18 basis point decrease in delinquency rates one year afterwards. All of these effects are equivalent to about 5 to 10 percent of the outcome standard deviation. In addition, unlike in the capital results, all coefficients on the CAMELS asset quality-specific ratings are highly significant and increase monotonically as the CAMELS ratings become worse.

Management

The regressions using the management text utilize the same outcome variables as the composite score and are reported in table 6. The coefficient of interest is negative as expected and is significant across all specifications. Positive sentiment on the text in the management section is associated with lower MRAs and MRIAs at the next comprehensive exam. Note that the magnitudes of the coefficients are much lower than in the composite regressions. This decreased explanatory power is consistent with MRAs/MRIAs not all necessarily being associated with management specific issues. For example, particular MRAs and MRIAs may relate to asset quality issues. The results, however, are still highly statistically and economically significant. As more issues are identified, it is more likely that management is an issue. Indeed, the CAMELS management-specific ratings are also highly significant and load monotonically in magnitude, as in the case of the composite ratings regressions (table 3).

Earnings

Table 7 shows the results of regressing measures of profitability on sentiment scores. The left hand side variables are return on assets (ROA) and pre-provision net revenue (PPNR) divided by assets. Higher values for both of these ratios are associated with higher earnings and profitability. The coefficient of interest on lag sentiment is positive and statistically significant across all specifications, indicating that higher sentiment is associated with improved profitability one year later. A one-standard deviation improvement in either the polar or valence LM sentiment existing in the Earnings section of bank exams (0.24 and 0.10, respectively) leads to about 6 to 8 basis point increase in ROA and PPNR rates. These results are both statistically and economically significant results. In fact, these effects are relatively substantial compared to the standard deviations of ROA and PPNR rates, which are 39 and 53 basis points, respectively.

Liquidity

Table 8 shows the results of regressing liquidity measures on sentiment scores. The first measure is securities as a share of assets. The second measure is cash and securities as a share of assets. The coefficient of interest changes signs and is not statistically different from zero. While these outcome measures are clearly related to liquidity, a measure of a bank's liquidity *needs* is not included. This latter characteristic is harder to determine from a bank's financial statements. Supervisors look at both liquidity levels and needs when assessing this CAMELS score. Since the financial crisis, internal liquidity stress tests are also reviewed. These data is not available in the Call Reports, however.[5] In addition, the word counts for this section are generally low, making it hard for the sentiment score to capture additional information. Also, note that the econometric models are able to explain the variation in the dependent variables. When it comes to projecting securities to assets or cash and securities to assets, year fixed effects, the CAMELS dummies, and the lag in these variables account for the vast majority of the variation. R-squareds for all specifications are greater than 0.86. Also, as seen in section 4 (table 1), we know that our method of calculating sentiment is not as highly correlated with the liquidity CAMELS component. Therefore, measurement error may be driving the insignificant results.

8 Results Using Subsamples and Interactions

In this section, we provide more results based on subsamples of the bank exam data and based on interactions of sentiment with other variables. In particular, we are interested in whether there are certain circumstances in which sentiment in the bank examination process is more helpful in providing insights to bank future outcomes. Additional regressions comparing banks by size are reported in the appendix section 10.2.

8.1 Strong Score Versus Weak Score

The next set of regressions relate bank outcomes to sentiment but separately test whether banks that had a strong or weak rating in the prior exam for that section of the exam matters in predicting bank outcomes. Recall from figure 2 that most banks have a strong score (1 or 2). When a bank has a lower quality score (3, 4, or 5), especially for the composite score, the bank is subject to more restrictions. As a result, banks have strong incentives to improve bank outcomes (and their scores) while examiners are likely focused on making sure they help prevent any further deterioration in the safety and soundness of banks. For MRAs and MRIAs, sufficient leeway may be important, especially for strong banks, such that banks are not surprised by any information revealed during the examination process. For most other metrics, the bank examination process may focus on the weak banks; signalling supervisory views may be especially important in returning banks to better safety and soundness conditions.

[5] We also tested liquidity measures that include types of funding (large time deposits/total liabilities, large time deposits/total assets, (cash + securities - large time)/total assets). These results are also not statistically significant and are not reported for conciseness.

The sentiment score regression is run separately depending on the CAMELS score. The specifications are as follows

$$\begin{aligned}\text{outcome}_{i,t} &= \rho_s \text{ outcome}_{t-1} + \beta_s \text{ sentiment}_{c,t-1} + \gamma_s \; log(\text{assets}_{t-1}) \\ &+ \Sigma_{n=1}^{4} \; \psi_n \text{ CAMELS dummy}_{i,c,n,t-1} + \phi_t + \epsilon_{i,t}, \text{if } t-1 \text{ component rating } \in [1,2] \text{ and} \\ \text{outcome}_{i,t} &= \rho_w \text{ outcome}_{t-1} + \beta_w \text{ sentiment}_{c,t-1} + \gamma_w \; log(\text{assets}_{t-1}) \\ &+ \Sigma_{n=1}^{4} \; \psi_n \text{ CAMELS dummy}_{i,c,n,t-1} + \phi_t + \epsilon_{i,t}, \text{if } t-1 \text{ component rating } \in [3,4,5],\end{aligned}$$

where the subscript s is for a strong bank and w is for a weak bank. The coefficient of interest is still β, and the test results are summarized in table 9.

We generally find that our results from the benchmark regressions appear to be driven by strong banks when it comes to the sentiment for overall and management sections in predicting MRAs and MRIAs (coefficients are more statistically significant, have smaller standard errors, and have larger magnitudes). In contrast, the predictive power embedded in other sections are driven by the weak bank samples. In particular, the sentiment in the capital adequacy sections seem to provide some predictability for future capital ratios for weak banks.

Going through each component, for the composite score that uses all of the exam text, the estimated β is negative and significant across all specifications. As sentiment increases, MRAs are likely to decrease regardless of the composite score. The coefficient magnitudes are very similar to the results in table 3 regardless of whether the sample is for weak or strong banks. When looking at capital adequacy, the coefficients on the sentiment scores are statistically significant for the weak banks using the polar sentiment. This is consistent with bank examiners describing improvements at weak banks that help them build back capital by the next exam. This is also consistent with our findings in table 4 that showed that there was quite a drop off in capital ratios when the capital adequacy rating was the lowest ratings in terms of capital ratios. However, the weak bank sample is relatively small. There are about 700 observations in the specification for columns 1 and 3 for capital. When it comes to asset quality, the β coefficients are still negative as in the baseline analysis and generally statistically significant. The results are much stronger for banks with weak asset quality scores both in terms of statistical significance and the magnitude of the coefficient. This result implies that supervisors tend to focus on conveying more value-added and relevant information about future bank outcomes through their write-ups, especially in times of weak asset quality at banks. The β coefficients are negative and statistically significant for the management regressions. In particular, the magnitudes of the coefficients are stronger for strong banks. This is consistent with supervisors perhaps feeling the need to express more information about management quality for strong banks when it comes to MRA/MRIAs as these are more prevalent (and less surprising) at weak banks. As with the regressions for the asset quality sections, the earnings results are much stronger for weak banks in magnitude. Both sets of regressions have statistical significance. The earnings results can also imply that supervisors tend to convey more information about earnings prospects at weak banks, similar to the sentiment embodied in the asset quality sections. Finally, the sentiment associated with the liquidity sections of the exams continue to be insignificant, indicating that the non-result applies to subsamples as well.

8.2 GFC Versus Post-GFC

The next set of regressions relate bank outcomes and sentiment but separately for exams during the financial crisis and after the financial crisis to see if supervisors conveyed information relevant to future outcomes differently in a particular period.

The sentiment score regression is run separately using the following specifications

$$\begin{aligned} \text{outcome}_t &= \rho_{GFC}\ \text{outcome}_{t-1} + \beta_{GFC}\ \text{sentiment}_{c,t-1} + \gamma_{GFC}\ log(\text{assets}_{t-1}) \\ &\quad + \Sigma_{n=1}^{4}\ \psi_n\ \text{CAMELS dummy}_{i,c,n,t-1} + \phi_t + \epsilon_{i,t}, \text{if } t \leq 2011 \text{ and} \\ \text{outcome}_t &= \rho_{post}\ \text{outcome}_{t-1} + \beta_{post}\ \text{sentiment}_{c,t-1} + \gamma_{post}\ log(\text{assets}_{t-1}) \\ &\quad + \Sigma_{n=1}^{4}\ \psi_n\ \text{CAMELS dummy}_{i,c,n,t-1} + \phi_t + \epsilon_{i,t}, \text{if } t \geq 2012, \end{aligned}$$

where the GFC period is defined between 2006 and 2011, and the *post* period is defined between 2012 and 2016:Q2. The coefficient of interest is still β. The results are summarized in table 10.

Similar to the regression results in table 9, the GFC period (up to 2011) is where most of the extra information embedded in the capital adequacy, asset quality, and earning sections were helpful in predicting future bank outcomes related to capital ratios, asset quality measures, and profitability.

More specifically, the β coefficients for the composite score regressions continue to be negative and statistically significant in all of the specifications; consistent with the weak-strong sample split, it appears that the coefficients are higher in magnitude for the post-GFC period when there were relatively more strong banks in the sample. The signs on the coefficients for capital adequacy are mixed. During the financial crisis, positive sentiment is associated with an increase in capital by the next exam. The opposite is true after the financial crisis. This is consistent with trends in the banking industry. The system needed more capital during the crisis. After the crisis, banks built up capital, and the stronger banks started decreasing capital as they navigated the new regulatory regime; a positive sentiment in this environment may have motivated banks to be more relaxed in this dimension. The β coefficients remain negative in the asset quality tests. The statistical significance and magnitudes are higher during the financial crisis, possibly due to greater variation during the GFC when there were relatively more banks in the sample with poor asset quality. Like the composite score, the management sentiment scores continue to be negative and statistically significant in all of the specifications, and they do not differ that much whether in the GFC or post-GFC period. The β coefficients on the earnings tests remain positive and significant across specifications. The magnitudes are smaller after the crisis. As with the asset quality results, this may be due to greater variation in earnings during the GFC when there were more weak banks in the sample. As with all the other regression results thus far, the liquidity sentiment measures generally are not statistically significant.

8.3 Interaction Between Sentiment and CAMELS Dummy

As another way to check to see if most of the extra informational content that predicts various bank outcomes come mostly from weak or strong banks for certain segments of the examination reports, this next set of regressions uses the full sample between 2004:Q1 and 2016:Q2 and tests the additional marginal impact of sentiment on the output variable by including an interactive term with the lagged dummy indicating if the bank had a weak score. The specifications follow this structure:

$$\begin{aligned}\text{outcome}_{i,t} \quad &= \quad \rho\ \text{outcome}_{i,t-1} + \beta\ \text{sentiment}_{i,c,t-1} + \delta\ \text{sentiment}_{i,c,t-1} * \text{weak dummy}_{i,c,t-1} \\ &\quad +\gamma\ log(\text{assets}_{i,t-1}) + \psi\ \text{weak dummy}_{i,c,t-1} + \phi_t + \epsilon_{i,t}.\end{aligned}$$

The coefficient δ on the interacted term provides the marginal effect between the sentiment score and the weak bank dummy variable.

Tables 11 and 12 each report the results for three of the types of outcome variables, focusing on the β and δ coefficients and comparing results to the baseline (columns 1 and 3). The results are consistent with the benchmark regressions and the subsample regressions. For composite and management regressions, we see that, in general, better sentiment is associated with fewer MRAs and MRIAs. However, for weak banks, the interaction terms mostly cancel out, the direct effects of better sentiment on (more) MRAs and MRIAs in weak banks offsets the effects on all banks. In contrast, for the asset quality and earnings regressions, the sentiment coefficient effect that sentiment has on future bank outcomes along asset quality and profitability measures, respectively, are accentuated for weak banks. Moreover, it is only for weak banks that better sentiment embedded in examination reports has a positive effect on future capital ratios.

Based on the subsample and interaction-based regression results, the evidence suggests that most of the predictive power of sentiment in bank examination reports are driven by strong banks when it comes to MRAs and MRIAs, and by weak banks when it comes to other bank outcomes such as capital adequacy, asset quality, and earnings. The liquidity regression results are generally insignificant statistically speaking and this may be due to the fact that the LM dictionary is not capturing sentiment well in this portion of the exam or the proxy for liquidity may not be the best.

8.4 Longer Term Effects

Finally, we explore whether information in bank exam text is associated with bank outcomes over a longer time period, namely two years or two exam cycles. The first set of regressions follow a similar specification as the baseline, but shifts the lagged right hand side variables back one more year:

$$
\begin{aligned}
\text{outcome}_{i,t} \quad = \quad & \rho \, \text{outcome}_{i,t-2} + \beta \, \text{sentiment}_{i,c,t-2} + \gamma \, log(\text{assets}_{i,t-2}) \\
& + \Sigma_{n=1}^{4} \, \psi_n \, \text{CAMELS dummy}_{i,c,n,t-2} + \phi_t + \epsilon_{i,t}.
\end{aligned}
$$

β remains the coefficient of interest.[6] For the MRA/MRIA variables, the outcome is measured as the average over the next two exam cycles. In other words, the regression specification is testing whether, on average, the number of MRA/MRIAs are lower for the two years after an exam.

The results are reported in table 13. The first two columns repeat the baseline results shown before, reporting only the beta coefficients on specifications testing one year after an exam. The next two columns report results testing for outcomes two years after the exam. All the results suggest that even for a longer horizon, our results are generally statistically significant, though the coefficients are understandably smaller in magnitude. This suggests that the sentiment conveyed in the bank examination reports has both shorter-term and longer-term insights into bank outcomes.

9 Conclusion

In this paper, we analyze whether the bank examination process provides useful insight into bank future outcomes by using textual analysis on commercial bank examination reports. In particular, we find that controlling for a variety of factors, the sentiment supervisors express in describing many of the components predict future bank outcomes. More specifically, the sentiment conveyed in the asset quality, management, and earnings sections provides significant information in predicting future outcomes for problem loans, supervisory actions, and profitability, respectively. We show that this relationship is driven by banks with better ratings when it comes to management, and

[6] Given that the liquidity sentiment measures never loaded in the prior tests, we drop them from this analysis.

banks with worse ratings when it comes to asset quality and earnings. In addition, the sentiment embodied in the capital adequacy section of the reports provide insights into future capital ratios only for weak banks. In fact bank examiners may have extra incentives to be precise with their language in these subcomponents as capital adequacy, asset quality, and earnings are critical to a bank as a going-concern. Overall, our results on the relationship between positive sentiment and future positive bank outcomes is most striking when it comes to MRA/MRIAs and bank earnings. All of this suggests that bank supervisors play a meaningful role in the surveillance of the banking system by creating and sharing information that is embedded in bank examination reports through the bank examination process.

There are several caveats to our analysis, however. First, we may be capturing the effects of other types of information in the bank examination process rather than the sentiment in the exams itself. However, even if this were true, this still implies that meaningful information is created, documented, and shared in the examination process. In turn, this is important for understanding the role that supervision plays in maintaining the safety and soundness of the banking system. Second, we only show that the examination process appears to help in monitoring future bank outcomes and the banking system as a whole. However, we have nothing to say about the efficiency of the process. For example, does the added-value of bank supervision outweigh the costs of maintaining a large number of personnel and resources in this process? Third, our analysis is largely for small to medium sized banks. The degree in which supervisory examination information is useful for bank outcomes for large bank holding companies may be different.

References

Bassett, William F., Seung Jung Lee and Thomas Popeck Spiller. 2015. "Estimating Changes in Supervisory Standards and Their Economic Effects." *Journal of Banking and Finance* 60:21–43.

Berger, Allen N. and Sally M. Davies. 1998. "The Information Content of Bank Examinations." *Journal of Financial Services Research* 14(2):117–144.

Berger, Allen N., Sally M. Davies and Mark J. Flannery. 2000. "Comparing Market and Supervisory Assessments of Bank Performance: Who Knows What When?" *Journal of Money, Credit and Banking* 32(3):641–667.

Correa, Ricardo, Keshav Garud, Juan Londono and Nathan Mislang. 2021. "Sentiment in Central Banks Financial Stability Reports." *Review of Finance* 25(1):85–120.

DeYoung, Robert, Mark J. Flannery, William W. Lang and Sorin M. Sorescu. 2001. "The Information Content of Bank Exam Ratings and Subordinated Debt Prices." *Journal of Money, Credit and Banking* 33(4):900–925.

Gandhi, P., T. Loughran and B. McDonald. 2019. "Using Annual Report Sentiment as a Proxy for Financial Distress in U.S. Banks." *Journal of Behavioral Finance* 20(4):424–436.

Gaul, Lewis and Jonathan Jones. 2021. "CAMELS Ratings and Their Information Content." *OCC Working Papers* .

Goldsmith-Pinkham, Paul, Beverly Hirtle and David Lucca. 2016. "Parsing the Content of Bank Supervision." *Federal Reserve Bank of New York Staff Reports* (770).

Gupta, Aparna, Cheng Lu, Majeed Simaan and Mohammed Zaki. 2022. "When Positive Sentiment is Not So Positive: Textual Analytics and Bank Failure." *Working Paper* .

Henry, Elaine. 2008. "Are Investors Influenced by How Earnings Press Releases Are Written?" *Journal of Business Communication* 45(4):363–407.

Hirtle, Beverly, Anna Kovner and Matthew Plosser. 2020. "The Implact of Supervision on Bank Performance." *Journal of Finance* 75(5):2765–2808.

Hu, Minqing and Bing Liu. 2004. "Mining Opinion Features in Customer Reviews." *proceedings of the tenth ACM SIGKDD International Conference on Knowledge Discovery and Data Mining* p. 168177.

Hubert, Paul and Fabien Labondance. 2017. "Central bank sentiment and policy expectations." *Bank of England Staff Working Paper* .

Jiang, Fuwei, Joshua Lee, Xiumin Martin and Guofu Zhou. 2019. "Manager sentiment and stock returns." *Journal of Financial Economics* 132(1):126 – 149.

Jordan, John S., Joe Peek and Eric S. Rosengren. 2000. "The Market Reaction to the Disclosure of Supervisory Actions: Implications for Bank Transparency." *Journal of Financial Intermediation* 9(3):298 – 319.

Kalamara, Eleni, Arthur Turrell, Chris Redl, George Kapetanios and Sujit Kapadia. 2020. "Making Text Count: Economic Forecasting Using Newspaper Text." *Bank of England Working Paper* .

Loughran, Tim and Bill McDonald. 2011. "When Is a Liability Not a Liability? Textual Analysis, Dictionaries, and 10-Ks." *Journal of Finance* 66(1):35–65.

Nopp, Clemens and Allan Hanbury. 2015. "Detecting Risks in the Banking System by Sentiment Analysis." *Working paper* .

Nyman, Rickard, Sujit Kapadia and David Tuckett. 2021. "News and Narratives in Financial Systems: Exploiting Big Data for Systemic Risk Assessment." *Journal of Economic Dynamics and Control* 127:104119.

Peek, Joe, Eric S. Rosengren and Geoffrey M. B. Tootell. 1999. "Is Bank Supervision Central to Central Banking?*." *The Quarterly Journal of Economics* 114(2):629–653.

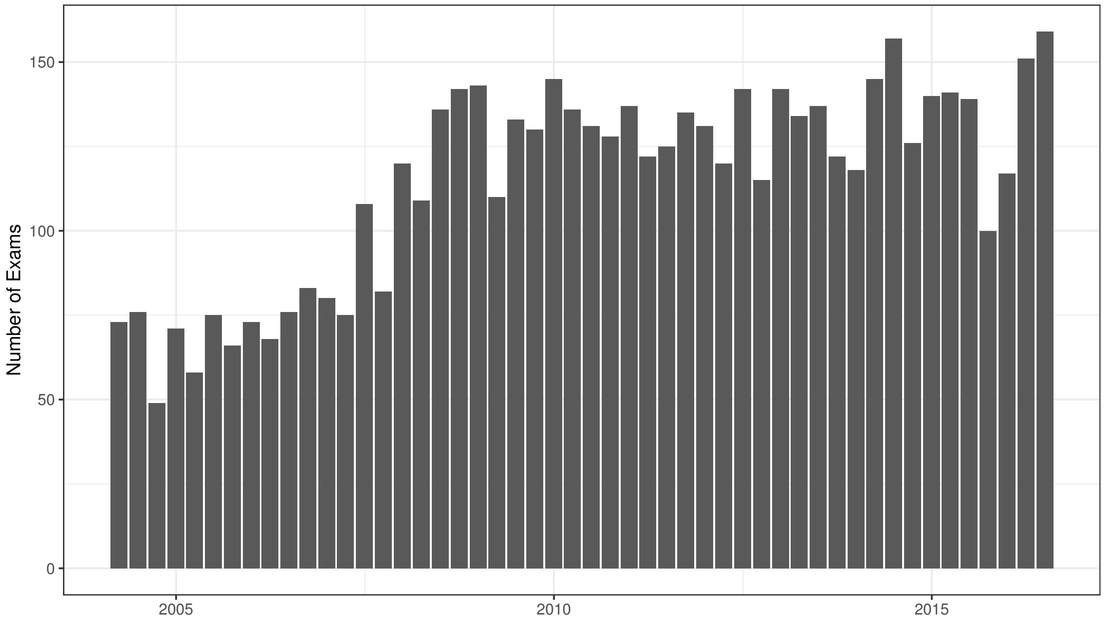
Price, S. McKay, James S. Doran, David R. Peterson and Barbara A. Bliss. 2012. "Earnings Conference Calls and Stock Returns: The Incremental Informativeness of Textual Tone." *Journal of Banking & Finance* 36(4):992 – 1011.

Rezende, Marcelo. 2014. "The Effects of Bank Charter Switching on Supervisory Ratings." *Finance and Economics Discussion Series* (2014-20).

Shapiro, Adam Hale and Daniel Wilson. 2019. "Taking the Fed at its Word: A New Approach to Estimating Central Bank Objectives using Text Analysis." *Federal Reserve Bank of San Francisco Working Paper* .

Shapiro, Adam Hale, Moritz Sudhof and Daniel Wilson. 2020. "Measuring News Sentiment." *Federal Reserve Bank of San Francisco Working Paper* .

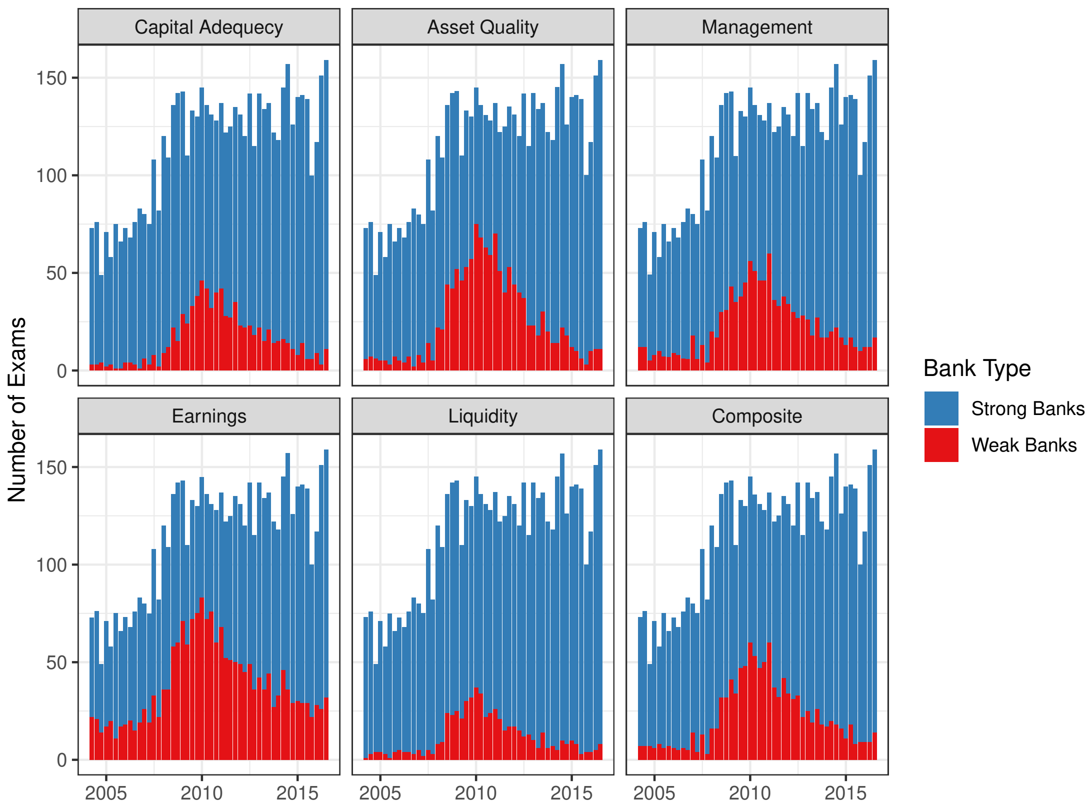
Figure 1: Exams in the Sample



Note: Data between 2004:Q1 and 2016:Q2.

Source: Confidential bank exams.

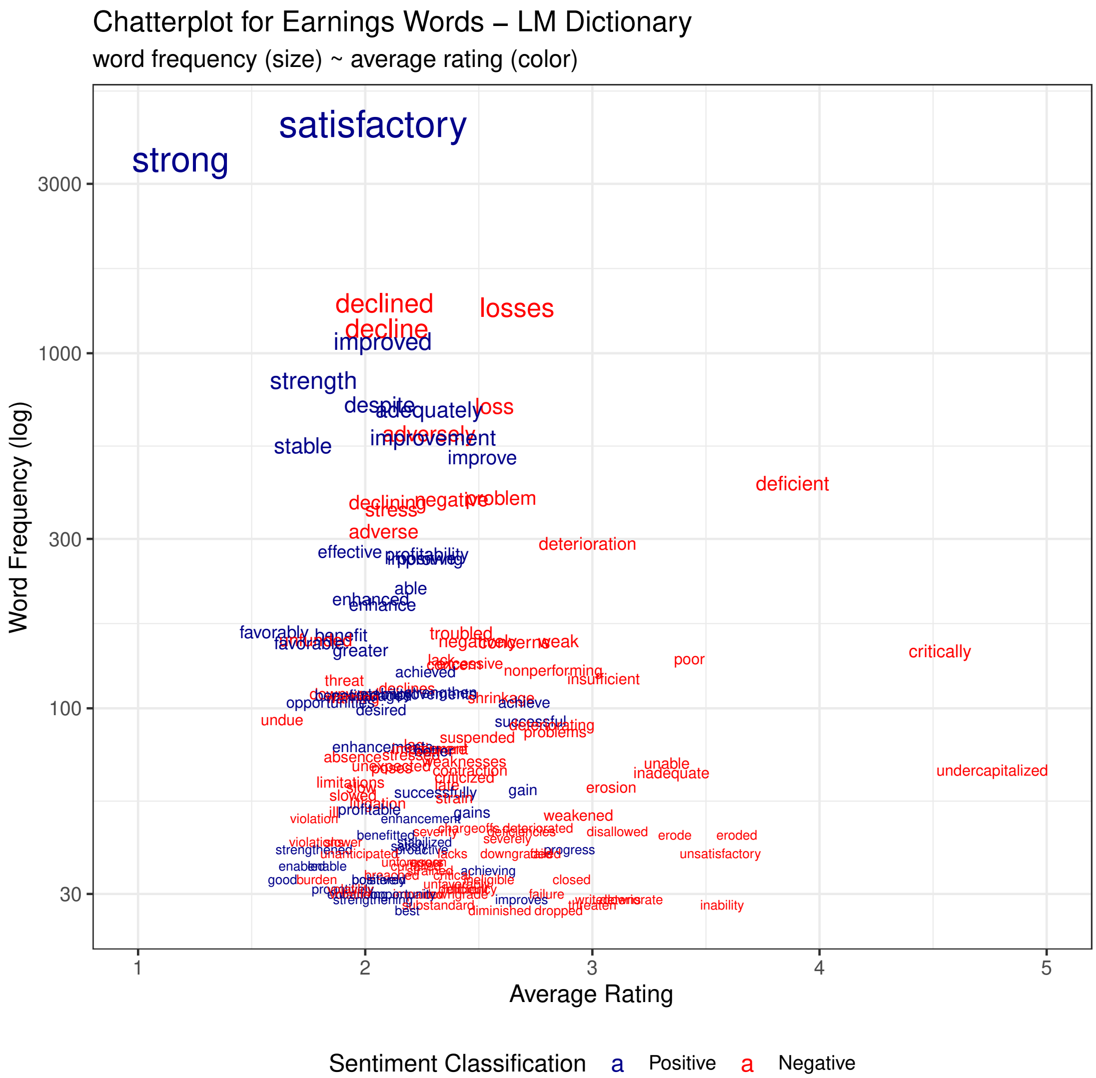
Figure 2: Distribution of CAMELS Component and Composite Scores



Note: Data between 2004:Q1 and 2016:Q2.

Source: Confidential bank exams.

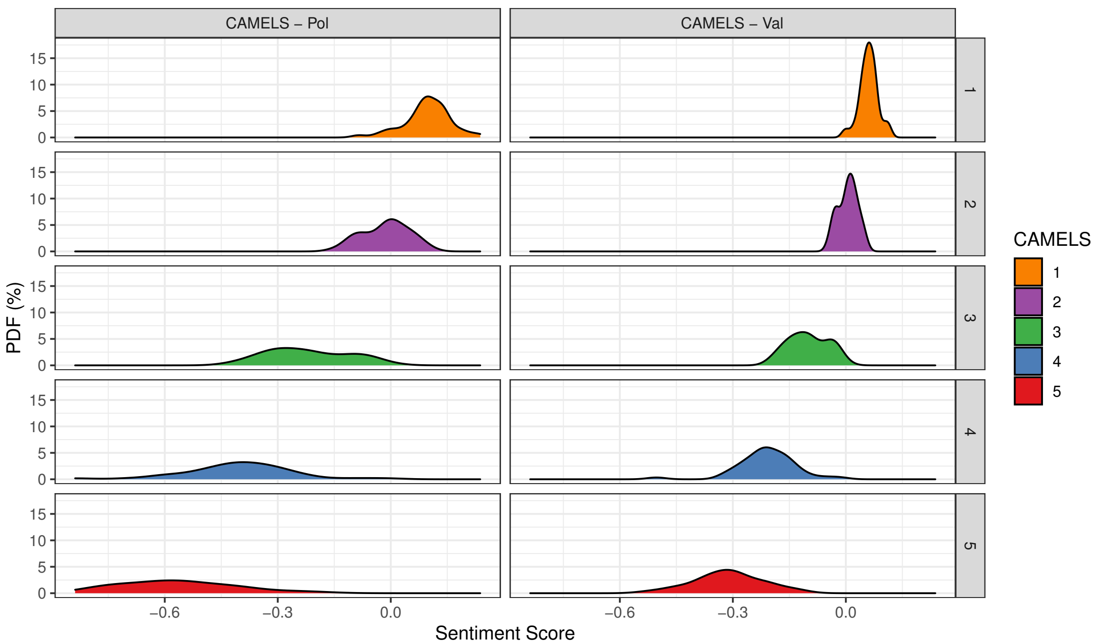
Figure 3: Frequencies and Average Ratings associated with Words in the LM Dictionary in Earnings Sections of Bank Exams



Note: Data between 2004:Q1 and 2016:Q2.

Source: Confidential bank exams.

Figure 4: LM Polar and Valence Sentiment Score Distributions by Composite CAMELS Score



Note: Data between 2004:Q1 and 2016:Q2. Distributions are kernel densities.

Source: Confidential bank exams.

Table 1: Sentiment Score Performance: Composite and Section CAMELS Score

This table reports the model performance (R-squared) from the specification

$$\text{CAMELS score}_s \quad = \quad \text{Sentiment score}_{s,m,l} + \epsilon_s,$$

where s is the section of the exam, m is the sentiment score method, and l is the sentiment score lexicon. The methods are listed in the first column: polar, polar (weighted, tf-idf), and valence. The second column is the bank exam section of text used, and the three lexicons are list across the top: FS, LM, and QDAP. The best performer based on R-squared or adjusted R-squared are listed in the last two columns.

	Exam	FS		LM		QDAP		Highest	
	section	R-squared	Adj. R-sqd.	R-squared	Adj. R-sqd.	R-squared	Adj. R-sqd.	R-squared	Adj. R-sqd.
Polar, regular	Total	0.16	0.16	0.25	0.25	0.10	0.10	LM	LM
	Capital	0.06	0.06	0.21	0.21	0.15	0.15	LM	LM
	Asset quality	0.03	0.03	0.10	0.10	0.08	0.08	LM	LM
	Management	0.13	0.13	0.14	0.14	0.12	0.12	LM	LM
	Earnings	0.20	0.20	0.30	0.30	0.32	0.32	QDAP	QDAP
	Liquidity	0.02	0.02	0.09	0.09	0.10	0.10	QDAP	QDAP
Polar, tf-idf	Total	0.15	0.14	0.10	0.10	0.14	0.14	FS	FS
	Capital	0.13	0.13	0.13	0.13	0.11	0.11	LM	LM
	Asset quality	0.05	0.05	0.07	0.06	0.05	0.05	LM	LM
	Management	0.15	0.15	0.10	0.10	0.13	0.13	FS	FS
	Earnings	0.23	0.23	0.24	0.24	0.25	0.25	QDAP	QDAP
	Liquidity	0.02	0.02	0.05	0.05	0.05	0.05	QDAP	QDAP
Valence	Total	0.20	0.20	0.33	0.33	0.04	0.04	LM	LM
	Capital	0.12	0.12	0.29	0.28	0.20	0.20	LM	LM
	Asset quality	0.08	0.08	0.18	0.18	0.12	0.12	LM	LM
	Management	0.14	0.14	0.18	0.18	0.16	0.16	LM	LM
	Earnings	0.24	0.24	0.35	0.35	0.37	0.37	QDAP	QDAP
	Liquidity	0.03	0.03	0.14	0.14	0.13	0.13	LM	LM

Table 2: Summary Statistics

This table reports summary statistics of variables used in the baseline regression tests. Control variables that are ratios have been winsorized at 1 and 99 percent. However, we show the 5th and 95th percentiles of all the variables to maintain confidentiality of the data.

Variable			Obs	Mean	Std. Dev.	5th Percent.	95th Percent.
Sentiment Scores							
	Polar						
		Composite	5,419	0.480	0.132	0.250	0.692
		Capital	5,500	0.637	0.273	0.182	1
		Asset quality	5,500	0.315	0.157	0.091	0.583
		Management	5,419	0.605	0.242	0.200	1
		Earnings	5,500	0.579	0.239	0.182	1
		Liquidity	5,500	0.610	0.246	0.222	1
	Valence						
		Composite	5,419	0.495	0.062	0.379	0.586
		Capital	5,500	0.536	0.085	0.381	0.657
		Asset quality	5,500	0.426	0.075	0.306	0.550
		Management	5,419	0.536	0.090	0.372	0.645
		Earnings	5,500	0.530	0.098	0.353	0.673
		Liquidity	5,500	0.534	0.071	0.410	0.645
Control Variables							
		MRIA/MRA sum	5,419	2.871	5.301	0	14
		MRIA/MRA dummy	5,419	0.387	0.487	0	1
		Tier 1 ratio	5,500	15.851	10.761	9.420	28.139
		CET1 ratio	5,495	15.765	10.381	9.371	28.155
		Loan loss provisions/loans	5,497	0.302	0.579	0	1.351
		4-qtr net charge-offs/assets	5,425	0.110	0.190	-0.010	0.487
		Delinquent loans/loans	5,497	2.639	2.651	0.078	7.908
		4-qtr ROA	5,500	0.310	0.391	-0.357	0.798
		4-qtr PPNR/assets	5,500	0.781	0.529	-0.044	1.523
		Securities/assets	5,500	21.062	13.787	2.368	47.864
		(Cash+securities)/assets	5,500	28.173	14.606	8.446	55.856
		CAMELS dummy 2	5,500	0.162	0.369	0	1
		CAMELS dummy 3	5,500	0.055	0.229	0	1
		CAMELS dummy 4	5,500	0.022	0.146	0	0
		CAMELS dummy 5	5,500	0.007	0.081	0	0
		Ln(assets)	5,500	12.369	1.348	10.369	14.874

Source: Call Reports and confidential bank exams.

Table 3: Composite Score Regressions

This table shows the regression results from testing

$$\begin{aligned} \text{outcome}_{i,t} &= \rho\, \text{outcome}_{i,t-1} + \beta\, \text{sentiment}_{i,t-1} + \gamma\, log(\text{assets}_{i,t-1}) \\ &+ \Sigma_{n=1}^{4}\, \psi_n\, \text{CAMELS dummy}_{i,n,t-1} + \phi_t + \epsilon_{i,t}, \end{aligned}$$

where outcome is either the summation of all matters requiring [immediate] attention (MRA/MRIA) for a bank or a dummy variable that takes the value of one if any MRAs or MRIAs exist for a bank and zero otherwise. The CAMELS dummy is based on the composite score. The sentiment score is based on all exam text.

	(1)	(2)	(3)	(4)
	MRA/MRIA Sum		MRA/MRIA Dummy	
VARIABLES	Polar	Valence	Polar	Valence
Lag sentiment	-8.404***	-15.99***	-0.756***	-1.473***
	(0.683)	(1.668)	(0.0495)	(0.112)
Lag MRA/MRIA sum	-0.155***	-0.162***		
	(0.0258)	(0.0257)		
Lag MRA/MRIA dummy			-0.410***	-0.424***
			(0.0184)	(0.0183)
Lag CAMELS 2 dummy	1.269***	1.317***	0.0985***	0.102***
	(0.122)	(0.126)	(0.0144)	(0.0146)
Lag CAMELS 3 dummy	3.838***	3.919***	0.224***	0.228***
	(0.346)	(0.354)	(0.0238)	(0.0243)
Lag CAMELS 4 dummy	4.309***	4.311***	0.166***	0.162***
	(0.697)	(0.717)	(0.0401)	(0.0409)
Lag CAMELS 5 dummy	4.239***	4.127***	0.182***	0.166***
	(1.010)	(1.028)	(0.0498)	(0.0505)
Lag ln(total assets)	0.0750	0.0645	0.0333***	0.0327***
	(0.0696)	(0.0704)	(0.00615)	(0.00620)
Constant	6.007***	10.00***	0.555***	0.932***
	(0.896)	(1.181)	(0.0827)	(0.103)
Observations	5,259	5,259	5,259	5,259
R-squared	0.282	0.275	0.456	0.450
Adj. R-squared	0.274	0.267	0.450	0.444
Fixed effects	year	year	year	year

Robust standard errors are in parentheses. *** p<0.01, ** p<0.05, * p<0.1

Table 4: Capital Score Regressions

This table shows the regression results from testing

$$\begin{aligned} \text{outcome}_{i,t} &= \rho\,\text{outcome}_{i,t-1} + \beta\,\text{sentiment}_{i,t-1} + \gamma\,log(\text{assets}_{i,t-1}) \\ &+\Sigma_{n=1}^{4}\,\psi_n\,\text{CAMELS dummy}_{i,n,t-1} + \phi_t + \epsilon_{i,t}, \end{aligned}$$

where outcome is either the tier 1 ratio or the common equity tier 1 (CET1) ratio. The CAMELS dummy is based on the capital score. The sentiment score is based on exam text in the capital section.

	(1)	(2)	(3)	(4)
	Tier 1 Ratio		CET1 Ratio	
VARIABLES	Polar	Valence	Polar	Valence
Lag sentiment	0.289	1.029	0.307	1.058
	(0.280)	(1.025)	(0.273)	(0.993)
Lag Tier 1 ratio	0.774***	0.774***		
	(0.0654)	(0.0655)		
Lag CET1 ratio			0.767***	0.767***
			(0.0614)	(0.0615)
Lag capital 2 dummy	-0.395**	-0.393**	-0.414**	-0.413**
	(0.186)	(0.183)	(0.184)	(0.181)
Lag capital 3 dummy	-0.103	-0.0934	-0.108	-0.100
	(0.226)	(0.218)	(0.215)	(0.208)
Lag capital 4 dummy	-0.428	-0.392	-0.445	-0.411
	(0.340)	(0.324)	(0.328)	(0.316)
Lag capital 5 dummy	-2.174***	-2.116***	-2.225***	-2.168***
	(0.411)	(0.385)	(0.391)	(0.371)
Lag ln(total assets)	-0.131*	-0.131*	-0.141**	-0.141**
	(0.0739)	(0.0737)	(0.0682)	(0.0680)
Constant	4.901***	4.528***	5.102***	4.726***
	(1.721)	(1.569)	(1.584)	(1.461)
Observations	5,500	5,500	5,495	5,495
R-squared	0.805	0.805	0.799	0.799
Adj. R-squared	0.803	0.803	0.797	0.797
Fixed effects	year	year	year	year

Robust standard errors are in parentheses. *** p<0.01, ** p<0.05, * p<0.1

Table 5: Asset Quality Score Regressions

This table shows the regression results from testing

$$\begin{aligned} \text{outcome}_{i,t} &= \rho\, \text{outcome}_{i,t-1} + \beta\, \text{sentiment}_{i,t-1} + \gamma\, log(\text{assets}_{i,t-1}) \\ &+ \Sigma_{n=1}^{4}\, \psi_n\, \text{CAMELS dummy}_{i,n,t-1} + \phi_t + \epsilon_{i,t}, \end{aligned}$$

where outcome is loan loss provisions/loans, net charge-offs/loans, or delinquencies/loans. The CAMELS dummy is based on the asset quality score. The sentiment score is based on exam text in the asset quality section.

	(1)	(2)	(3)	(4)	(5)	(6)
	Loan Loss Provisions/Loans		4-qtr Net Charge-offs/Loans		Delinquencies/Loans	
VARIABLES	Polar	Valence	Polar	Valence	Polar	Valence
Lag sentiment	-0.247***	-0.707***	-0.0895***	-0.260***	-0.808***	-2.357***
	(0.0362)	(0.0943)	(0.0139)	(0.0343)	(0.153)	(0.369)
Lag loan loss provisions/loans	0.336***	0.330***				
	(0.0255)	(0.0255)				
Lag 4-qtr net charge-offs/loans			0.407***	0.396***		
			(0.0259)	(0.0262)		
Lag delinquent loans/loans					0.716***	0.709***
					(0.0183)	(0.0184)
Lag assets 2 dummy	-0.0450**	-0.0474**	-0.0291***	-0.0300***	-0.377***	-0.384***
	(0.0197)	(0.0196)	(0.00595)	(0.00594)	(0.0672)	(0.0676)
Lag assets 3 dummy	0.138***	0.130***	0.0599***	0.0570***	0.201	0.179
	(0.0354)	(0.0356)	(0.0120)	(0.0121)	(0.125)	(0.126)
Lag assets 4 dummy	0.326***	0.306***	0.173***	0.167***	1.160***	1.104***
	(0.0791)	(0.0793)	(0.0280)	(0.0280)	(0.256)	(0.255)
Lag assets 5 dummy	0.752***	0.730***	0.372***	0.365***	3.154***	3.103***
	(0.167)	(0.166)	(0.0496)	(0.0496)	(0.605)	(0.603)
Lag ln(total assets)	0.0405***	0.0395***	0.0143***	0.0140***	0.0470**	0.0425**
	(0.00561)	(0.00565)	(0.00207)	(0.00208)	(0.0198)	(0.0199)
Constant	-0.261***	-0.0222	-0.0892***	-0.000815	0.426*	1.251***
	(0.0694)	(0.0846)	(0.0255)	(0.0303)	(0.255)	(0.305)
Observations	5,497	5,497	5,310	5,310	5,497	5,497
R-squared	0.416	0.419	0.407	0.411	0.612	0.613
Adj. R-squared	0.410	0.413	0.401	0.405	0.608	0.609
Fixed effects	year	year	year	year	year	year

Robust standard errors are in parentheses. *** p<0.01, ** p<0.05, * p<0.1

Table 6: Management Score Regressions

This table shows the regression results from testing

$$\begin{aligned} \text{outcome}_{i,t} &= \rho \text{ outcome}_{i,t-1} + \beta \text{ sentiment}_{i,t-1} + \gamma \ log(\text{assets}_{i,t-1}) \\ &+\Sigma_{n=1}^{4} \psi_n \text{ CAMELS dummy}_{i,n,t-1} + \phi_t + \epsilon_{i,t}, \end{aligned}$$

where outcome is either the summation of all matters requiring [immediate] attention (MRA/MRIA) for a bank or a dummy variable that takes the value of one if any MRAs or MRIAs exist for a bank and zero otherwise. The CAMELS dummy is based on the management score. The sentiment score is based on exam text in the management section.

	(1)	(2)	(3)	(4)
	MRA/MRIA Sum		MRA/MRIA Dummy	
VARIABLES	Polar	Valence	Polar	Valence
Lag sentiment	-2.192***	-5.521***	-0.258***	-0.623***
	(0.319)	(0.960)	(0.0242)	(0.0671)
Lag MRA/MRIA Sum	-0.163***	-0.165***		
	(0.0262)	(0.0262)		
Lag MRA/MRIA dummy			-0.426***	-0.430***
			(0.0186)	(0.0185)
Lag management 2 dummy	1.881***	1.920***	0.154***	0.159***
	(0.119)	(0.118)	(0.0146)	(0.0145)
Lag management 3 dummy	4.812***	4.837***	0.277***	0.282***
	(0.310)	(0.312)	(0.0220)	(0.0222)
Lag management 4 dummy	6.222***	6.172***	0.291***	0.289***
	(0.675)	(0.680)	(0.0385)	(0.0388)
Lag management 5 dummy	8.255***	8.115***	0.344***	0.332***
	(1.331)	(1.352)	(0.0560)	(0.0569)
Lag ln(total assets)	0.164**	0.166**	0.0402***	0.0405***
	(0.0725)	(0.0725)	(0.00621)	(0.00623)
Constant	1.573*	3.167***	0.220***	0.393***
	(0.905)	(1.000)	(0.0781)	(0.0865)
Observations	5,259	5,259	5,259	5,259
R-squared	0.283	0.282	0.445	0.442
Adj. R-squared	0.275	0.273	0.439	0.436
Fixed effects	year	year	year	year

Robust standard errors are in parentheses. *** p<0.01, ** p<0.05, * p<0.1

Table 7: Earnings Score Regressions

This table shows the regression results from testing

$$\begin{aligned} \text{outcome}_{i,t} &= \rho \, \text{outcome}_{i,t-1} + \beta \, \text{sentiment}_{i,t-1} + \gamma \, log(\text{assets}_{i,t-1}) \\ &+ \Sigma_{n=1}^{4} \, \psi_n \, \text{CAMELS dummy}_{i,n,t-1} + \phi_t + \epsilon_{i,t}, \end{aligned}$$

where outcome is either ROA or PPNR/assets. The CAMELS dummy is based on the earnings score. The sentiment score is based on exam text in the earnings section.

	(1)	(2)	(3)	(4)
	Weighted 4-qtr ROA		Weighted 4-qtr PPNR/Assets	
VARIABLES	Polar	Valence	Polar	Valence
Lag sentiment	0.271***	0.741***	0.326***	0.861***
	(0.0199)	(0.0555)	(0.0263)	(0.0711)
Lag weighted 4-qtr ROA	0.595***	0.581***		
	(0.0202)	(0.0205)		
Lag weighted 4-qtr PPNR/assets			0.695***	0.690***
			(0.0239)	(0.0240)
Lag earnings 2 dummy	0.0148	0.0130	0.0172	0.0153
	(0.00935)	(0.00930)	(0.0131)	(0.0130)
Lag earnings 3 dummy	-0.0168	-0.0188	-0.0480**	-0.0496***
	(0.0158)	(0.0157)	(0.0192)	(0.0192)
Lag earnings 4 dummy	-0.0992***	-0.0910***	-0.0905***	-0.0792***
	(0.0281)	(0.0282)	(0.0260)	(0.0262)
Lag earnings 5 dummy	-0.411***	-0.391***	-0.289***	-0.262***
	(0.0489)	(0.0493)	(0.0374)	(0.0379)
Lag ln(total assets)	-0.000182	0.000158	0.0258***	0.0264***
	(0.00280)	(0.00281)	(0.00406)	(0.00404)
Constant	-0.0193	-0.255***	-0.273***	-0.545***
	(0.0342)	(0.0419)	(0.0462)	(0.0578)
Observations	5,501	5,501	5,501	5,501
R-squared	0.596	0.600	0.664	0.666
Adj. R-squared	0.592	0.596	0.661	0.663
Fixed effects	year	year	year	year

Robust standard errors are in parentheses. *** p<0.01, ** p<0.05, * p<0.1

Table 8: Liquidity Score Regressions

This table shows the regression results from testing

$$\begin{aligned}\text{outcome}_{i,t} &= \rho\,\text{outcome}_{i,t-1} + \beta\,\text{sentiment}_{i,t-1} + \gamma\,log(\text{assets}_{i,t-1})\\ &+\Sigma_{n=1}^{4}\,\psi_n\,\text{CAMELS dummy}_{i,n,t-1} + \phi_t + \epsilon_{i,t},\end{aligned}$$

where outcome is either securities/assets or (cash+securities)/assets. The CAMELS dummy is based on the liquidity score. The sentiment score is based on exam text in the liquidity section.

	(1)	(2)	(3)	(4)
	Securities/Assets		(Cash+Securities)/Assets	
VARIABLES	Polar	Valence	Polar	Valence
Lag sentiment	0.399	1.754*	-0.335	-1.864
	(0.273)	(0.964)	(0.325)	(1.135)
Lag securities/assets	0.941***	0.941***		
	(0.00582)	(0.00585)		
Lag cash+total securities/total assets			0.922***	0.923***
			(0.00720)	(0.00721)
Lag liquidity 2 dummy	-0.327*	-0.320*	-0.436**	-0.446**
	(0.178)	(0.178)	(0.213)	(0.213)
Lag liquidity 3 dummy	-0.633*	-0.603	-0.767*	-0.812*
	(0.371)	(0.371)	(0.451)	(0.454)
Lag liquidity 4 dummy	0.0743	0.127	-0.0977	-0.176
	(0.449)	(0.447)	(0.881)	(0.884)
Lag liquidity 5 dummy	0.927	1.043	0.280	0.126
	(1.203)	(1.216)	(1.267)	(1.257)
Lag ln(total assets)	-0.00243	-0.000825	-0.222***	-0.223***
	(0.0534)	(0.0534)	(0.0646)	(0.0645)
Constant	0.953	0.245	5.415***	6.214***
	(0.668)	(0.834)	(0.860)	(1.032)
Observations	5,500	5,500	5,500	5,500
R-squared	0.897	0.897	0.870	0.870
Adj. R-squared	0.896	0.896	0.869	0.869
Fixed effects	year	year	year	year

Robust standard errors are in parentheses. *** p<0.01, ** p<0.05, * p<0.1

Table 9: Regressions–Strong/Weak Separately

This table shows the coefficients and standard errors on the lagged sentiment score. Banks are classified as "strong" or "weak" based on rating in the prior exam for that section of the exam. Most banks have a strong score (1 or 2). Weak is defined as 3, 4, or 5. The full specifications are

$$\begin{aligned}
\text{outcome}_{i,t} &= \rho_s \text{ outcome}_{t-1} + \beta_s \text{ sentiment}_{t-1} + \gamma_s \; log(\text{assets}_{t-1}) \\
&\quad + \Sigma_{n=1}^{4} \; \psi_n \text{ CAMELS dummy}_{i,n,t-1} + \phi_t + \epsilon_{i,t}, \text{if } t-1 \text{ component rating } \in [1,2] \text{ and} \\
\text{outcome}_{i,t} &= \rho_w \text{ outcome}_{t-1} + \beta_w \text{ sentiment}_{t-1} + \gamma_w \; log(\text{assets}_{t-1}) \\
&\quad + \Sigma_{n=1}^{4} \; \psi_n \text{ CAMELS dummy}_{i,n,t-1} + \phi_t + \epsilon_{i,t}, \text{if } t-1 \text{ component rating } \in [3,4,5],
\end{aligned}$$

where the subscript s is for a strong bank and w is for a weak bank. The CAMELS dummy and the sentiment score are based on the relevant section.

		(1)	(2)	(3)	(4)
		Polar		Valence	
		Weak	Strong	Weak	Strong
Composite					
	MRA/MRIA Sum	-8.931***	-9.704***	-14.28***	-20.93***
		(2.184)	(0.726)	(4.118)	(1.819)
	MRA/MRIA Dummy	-0.559***	-0.713***	-1.007***	-1.440***
		(0.141)	(0.0536)	(0.258)	(0.127)
Capital					
	Tier 1 Ratio	0.823**	0.197	1.618	0.709
		(0.388)	(0.282)	(1.019)	(1.145)
	CET1 Ratio	0.851**	0.205	1.676	0.693
		(0.400)	(0.278)	(1.043)	(1.125)
Asset Quality					
	Loan Loss Provisions/Loans	-0.666***	-0.0411	-1.450***	-0.121*
		(0.148)	(0.0270)	(0.308)	(0.0625)
	4-qtr Net Charge-offs/Loans	-0.173***	-0.0295***	-0.420***	-0.0842***
		(0.0646)	(0.0104)	(0.119)	(0.0234)
	Delinquent Loans/Loans	-1.845***	-0.306**	-4.480***	-0.870***
		(0.663)	(0.127)	(1.252)	(0.302)
Management					
	MRA/MRIA Sum	-1.991*	-2.635***	-3.607	-7.609***
		(1.199)	(0.319)	(2.596)	(0.988)
	MRA/MRIA Dummy	-0.119*	-0.241***	-0.273*	-0.643***
		(0.0712)	(0.0257)	(0.144)	(0.0773)
Earnings					
	Weighted 4-qtr ROA	0.452***	0.0839***	1.053***	0.228***
		(0.0483)	(0.0135)	(0.110)	(0.0380)
	Weighted 4-qtr PPNR/Assets	0.464***	0.127***	1.067***	0.317***
		(0.0541)	(0.0199)	(0.121)	(0.0560)
Liquidity					
	Securities/Assets	0.212	0.365	0.822	1.746*
		(1.160)	(0.291)	(3.583)	(1.046)
	(Cash+Securities)/Assets	-2.731**	0.101	-7.503*	-0.0273
		(1.349)	(0.346)	(4.064)	(1.231)

Robust standard errors are in parentheses. *** p<0.01, ** p<0.05, * p<0.1

Table 10: Regressions–GFC and Post-GFC Separately

This table shows the coefficients and standard errors on the lagged sentiment score. GFC is defined as 2006–2011. Post GFC is defined as 2012–2016:Q2. The sentiment score regression is run separately using the following specifications

$$\begin{aligned} \text{outcome}_t &= \rho_{GFC}\ \text{outcome}_{t-1} + \beta_{GFC}\ \text{sentiment}_{t-1} + \gamma_{GFC}\ log(\text{assets}_{t-1}) \\ &\quad + \Sigma_{n=1}^{4}\ \psi_n\ \text{CAMELS dummy}_{i,n,t-1} + \phi_t + \epsilon_{i,t}, \text{if } t \geq 2012 \text{ and} \\ \text{outcome}_t &= \rho_{post}\ \text{outcome}_{t-1} + \beta_{post}\ \text{sentiment}_{t-1} + \gamma_{post}\ log(\text{assets}_{t-1}) \\ &\quad + \Sigma_{n=1}^{4}\ \psi_n\ \text{CAMELS dummy}_{i,n,t-1} + \phi_t + \epsilon_{i,t}, \text{if } t \leq 2011. \end{aligned}$$

The CAMELS dummy and the sentiment score are based on the relevant section.

		(1)	(2)	(3)	(4)
		Polar		Valence	
		GFC	Post GFC	GFC	Post GFC
Composite					
	MRA/MRIA Sum	-7.822***	-9.099***	-14.11***	-18.75***
		(0.928)	(0.767)	(2.177)	(1.885)
	MRA/MRIA Dummy	-0.664***	-0.856***	-1.257***	-1.754***
		(0.0582)	(0.0850)	(0.127)	(0.194)
Capital					
	Tier 1 Ratio	0.858**	-0.607*	2.568*	-1.491*
		(0.351)	(0.358)	(1.308)	(0.857)
	CET1 Ratio	0.868**	-0.579*	2.566**	-1.402*
		(0.344)	(0.345)	(1.276)	(0.834)
Asset Quality					
	Loan Loss Provisions/Loans	-0.353***	-0.0547*	-1.060***	-0.0899
		(0.0556)	(0.0301)	(0.142)	(0.0719)
	4-qtr Net Charge-offs/Loans	-0.114***	-0.0422***	-0.353***	-0.103***
		(0.0207)	(0.0143)	(0.0516)	(0.0312)
	Delinquent Loans/Loans	-1.099***	-0.231	-3.220***	-0.681
		(0.213)	(0.193)	(0.509)	(0.454)
Management					
	MRA/MRIA Sum	-2.159***	-2.164***	-4.981***	-6.080***
		(0.425)	(0.387)	(1.247)	(1.125)
	MRA/MRIA Dummy	-0.245***	-0.266***	-0.593***	-0.634***
		(0.0279)	(0.0384)	(0.0758)	(0.113)
Earnings					
	Weighted 4-qtr ROA	0.368***	0.115***	0.967***	0.320***
		(0.0291)	(0.0209)	(0.0775)	(0.0620)
	Weighted 4-qtr PPNR/Assets	0.431***	0.145***	1.110***	0.371***
		(0.0384)	(0.0221)	(0.101)	(0.0602)
Liquidity					
	Securities/Assets	0.636	0.175	2.755*	0.629
		(0.426)	(0.303)	(1.445)	(1.120)
	(Cash+Securities)/Assets	-0.513	0.0181	-2.305	-0.851
		(0.477)	(0.398)	(1.624)	(1.414)

Robust standard errors are in parentheses. *** p<0.01, ** p<0.05, * p<0.1

Table 11: Regressions–Interaction Between Sentiment and CAMELS Variable

This table shows the coefficients and standard errors on the lagged sentiment score. The specifications follow this structure:

$$\begin{aligned}\text{outcome}_{i,t} &= \rho\ \text{outcome}_{i,t-1} + \beta\ \text{sentiment}_{i,c,t-1} + \delta\ \text{sentiment}_{i,c,t-1} * \text{weak dummy}_{i,c,t-1}\\ &\quad +\gamma\ log(\text{assets}_{i,t-1}) + \psi\ \text{weak dummy}_{i,c,t-1} + \phi_t + \epsilon_{i,t}.\end{aligned}$$

The CAMELS dummy and the sentiment score are based on the relevant section.

		(1)	(2)	(3)	(4)
		Polar		Valence	
		Baseline	Interaction	Baseline	Interaction
Composite					
	MRA/MRIA Sum	-8.404***	-12.63***	-15.99***	-27.74***
		(0.683)	(0.709)	(1.668)	(1.757)
	Weak CAMELS Interaction		2.189		8.875**
			(2.072)		(4.011)
	MRA/MRIA Dummy	-0.756***	-0.910***	-1.473***	-1.887***
		(0.0495)	(0.0522)	(0.112)	(0.122)
	Weak CAMELS Interaction		0.770***		1.578***
			(0.152)		(0.287)
Capital					
	Tier 1 Ratio	0.289	0.295	1.029	1.084
		(0.280)	(0.292)	(1.025)	(1.194)
	Weak CAMELS Interaction		1.084**		2.336*
			(0.429)		(1.349)
	CET1 Ratio	0.307	0.307	1.058	1.081
		(0.273)	(0.287)	(0.993)	(1.170)
	Weak CAMELS Interaction		1.104**		2.408*
			(0.432)		(1.351)
Asset Quality					
	Loan Loss Provisions/Loans	-0.247***	-0.0367	-0.707***	-0.111*
		(0.0362)	(0.0285)	(0.0943)	(0.0657)
	Weak CAMELS Interaction		-0.937***		-1.928***
			(0.163)		(0.330)
	4-qtr Net Charge-offs/Loans	-0.0895***	-0.0234**	-0.260***	-0.0668***
		(0.0139)	(0.0109)	(0.0343)	(0.0247)
	Weak CAMELS Interaction		-0.286***		-0.627***
			(0.0668)		(0.123)
	Delinquent Loans/Loans	-0.808***	-0.254**	-2.357***	-0.732**
		(0.153)	(0.128)	(0.369)	(0.300)
	Weak CAMELS Interaction		-2.783***		-6.083***
			(0.661)		(1.266)

Robust standard errors are in parentheses. *** p<0.01, ** p<0.05, * p<0.1

Table 12: Regressions–Interaction Between Sentiment and CAMELS Variable (Cont.)

This table shows the coefficients and standard errors on the lagged sentiment score. The specifications follow this structure:

$$\begin{aligned} \text{outcome}_{i,t} &= \rho \text{ outcome}_{i,t-1} + \beta \text{ sentiment}_{i,c,t-1} + \delta \text{ sentiment}_{i,c,t-1} * \text{weak dummy}_{i,c,t-1} \\ &\quad + \gamma \, log(\text{assets}_{i,t-1}) + \psi \text{ weak dummy}_{i,c,t-1} + \phi_t + \epsilon_{i,t}. \end{aligned}$$

The CAMELS dummy and the sentiment score are based on the relevant section.

		(1)	(2)	(3)	(4)
		Polar		Valence	
		Baseline	Interaction	Baseline	Interaction
Management					
	MRA/MRIA Sum	-2.192***	-3.681***	-5.521***	-10.91***
		(0.319)	(0.331)	(0.960)	(1.021)
	Weak CAMELS Interaction		-0.320		2.608
			(1.202)		(2.617)
	MRA/MRIA Dummy	-0.258***	-0.317***	-0.623***	-0.868***
		(0.0242)	(0.0259)	(0.0671)	(0.0776)
	Weak CAMELS Interaction		0.308***		0.784***
			(0.0780)		(0.166)
Earnings					
	Weighted 4-qtr ROA	0.271***	0.115***	0.741***	0.322***
		(0.0199)	(0.0149)	(0.0555)	(0.0412)
	Weak CAMELS Interaction		0.437***		0.904***
			(0.0502)		(0.111)
	Weighted 4-qtr PPNR/Assets	0.326***	0.162***	0.861***	0.416***
		(0.0263)	(0.0198)	(0.0711)	(0.0551)
	Weak CAMELS Interaction		0.367***		0.746***
			(0.0545)		(0.122)
Liquidity					
	Securities/Assets	0.399	0.404	1.754*	1.958*
		(0.273)	(0.289)	(0.964)	(1.039)
	Weak CAMELS Interaction		-0.386		-2.358
			(1.163)		(3.553)
	(Cash+Securities)/Assets	-0.335	0.149	-1.864	0.241
		(0.325)	(0.347)	(1.135)	(1.230)
	Weak CAMELS Interaction		-2.441*		-6.799*
			(1.372)		(3.986)

Robust standard errors are in parentheses. *** p<0.01, ** p<0.05, * p<0.1

Table 13: Regressions–One and Two Years After an Exam

Columns 1 and 2 of this table repeat the results from the baseline results reported in tables 3 through 8, reporting the coefficients and standard errors on the lagged sentiment score. Columns 3 and 4 of this table show the coefficients and standard errors on the *twice* lagged sentiment score from the specification

$$\begin{aligned} \text{outcome}_{i,t} &= \rho \text{ outcome}_{i,t-2} + \beta \text{ sentiment}_{i,c,t-2} + \gamma \; log(\text{assets}_{i,t-2}) \\ &+ \Sigma_{n=1}^{4} \psi_n \text{ CAMELS dummy}_{i,n,t-2} + \phi_t + \epsilon_{i,t}. \end{aligned}$$

The CAMELS dummy and the sentiment score are based on the relevant section.

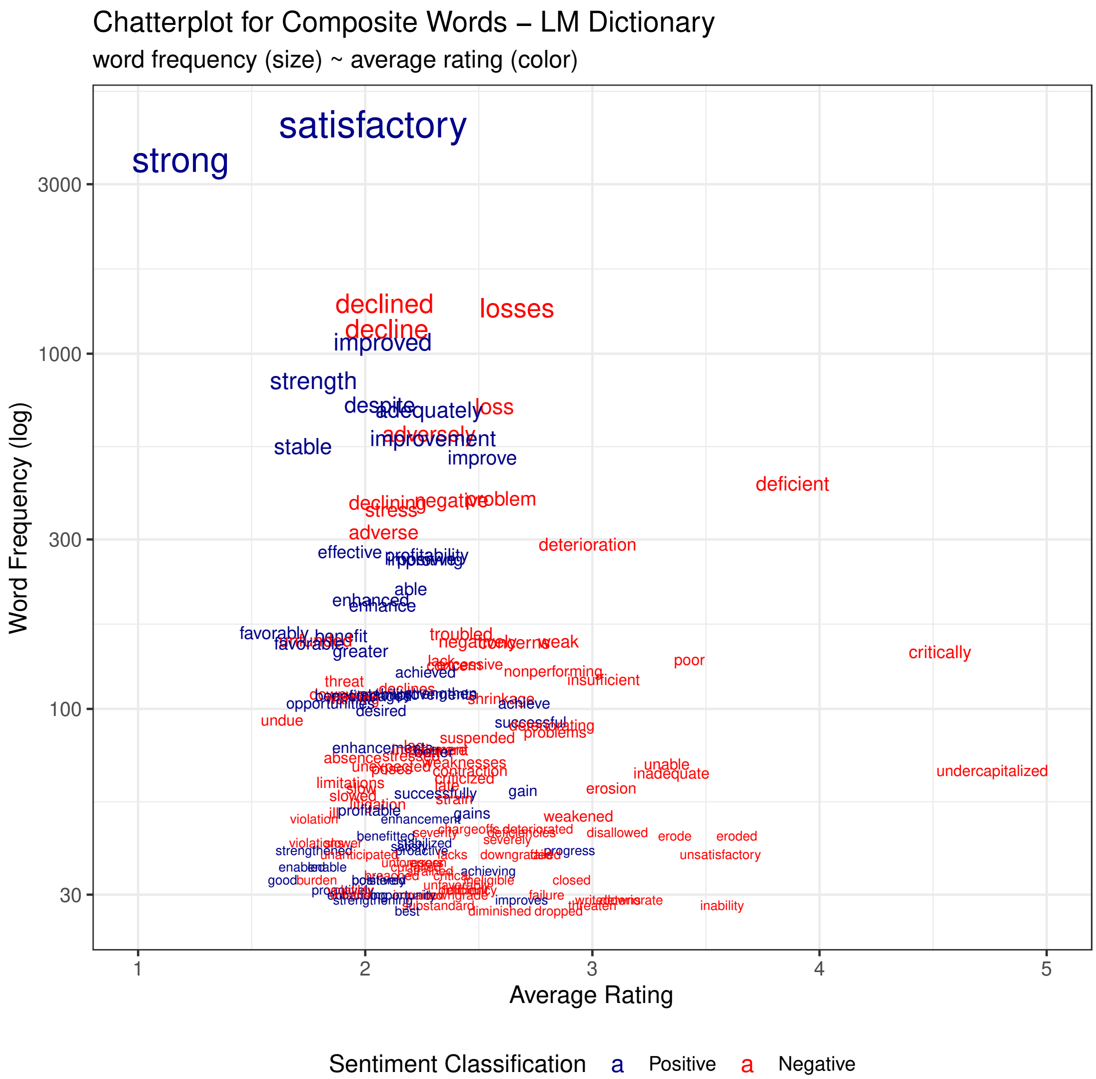
		(1)	(2)	(3)	(4)
		1 Year Out		2 Years Out	
		Polar	Valence	Polar	Valence
Composite					
	Average MRA/MRIA Sum	-8.404***	-15.99***	-2.971***	-5.734***
		(0.683)	(1.668)	(0.496)	(1.147)
	Average MRA/MRIA Dummy	-0.756***	-1.473***	-0.192***	-0.447***
		(0.0495)	(0.112)	(0.0294)	(0.0669)
Capital					
	Tier 1 Ratio	0.289	1.029	0.185	0.496
		(0.280)	(1.025)	(0.363)	(1.335)
	CET1 Ratio	0.307	1.058	0.212	0.576
		(0.273)	(0.993)	(0.351)	(1.284)
Asset Quality					
	Loan Loss Provisions/Loans	-0.247***	-0.707***	-0.105***	-0.358***
		(0.0362)	(0.0943)	(0.0398)	(0.0972)
	4-qtr Net Charge-offs/Loans	-0.0895***	-0.260***	-0.0657***	-0.188***
		(0.0139)	(0.0343)	(0.0140)	(0.0347)
	Delinquency Rate	-0.808***	-2.357***	-0.784***	-2.085***
		(0.153)	(0.369)	(0.181)	(0.426)
Management					
	Average MRA/MRIA Sum	-2.192***	-5.521***	-0.710***	-1.820***
		(0.319)	(0.960)	(0.192)	(0.560)
	Average MRA/MRIA Dummy	-0.258***	-0.623***	-0.0790***	-0.230***
		(0.0242)	(0.0671)	(0.0139)	(0.0396)
Earnings					
	Weighted 4-qtr ROA	0.271***	0.741***	0.225***	0.593***
		(0.0199)	(0.0555)	(0.0213)	(0.0584)
	Weighted 4-qtr PPNR/Assets	0.326***	0.861***	0.326***	0.850***
		(0.0263)	(0.0711)	(0.0298)	(0.0771)

Robust standard errors are in parentheses. *** p<0.01, ** p<0.05, * p<0.1

10 Appendix

10.1 More Chatterplots

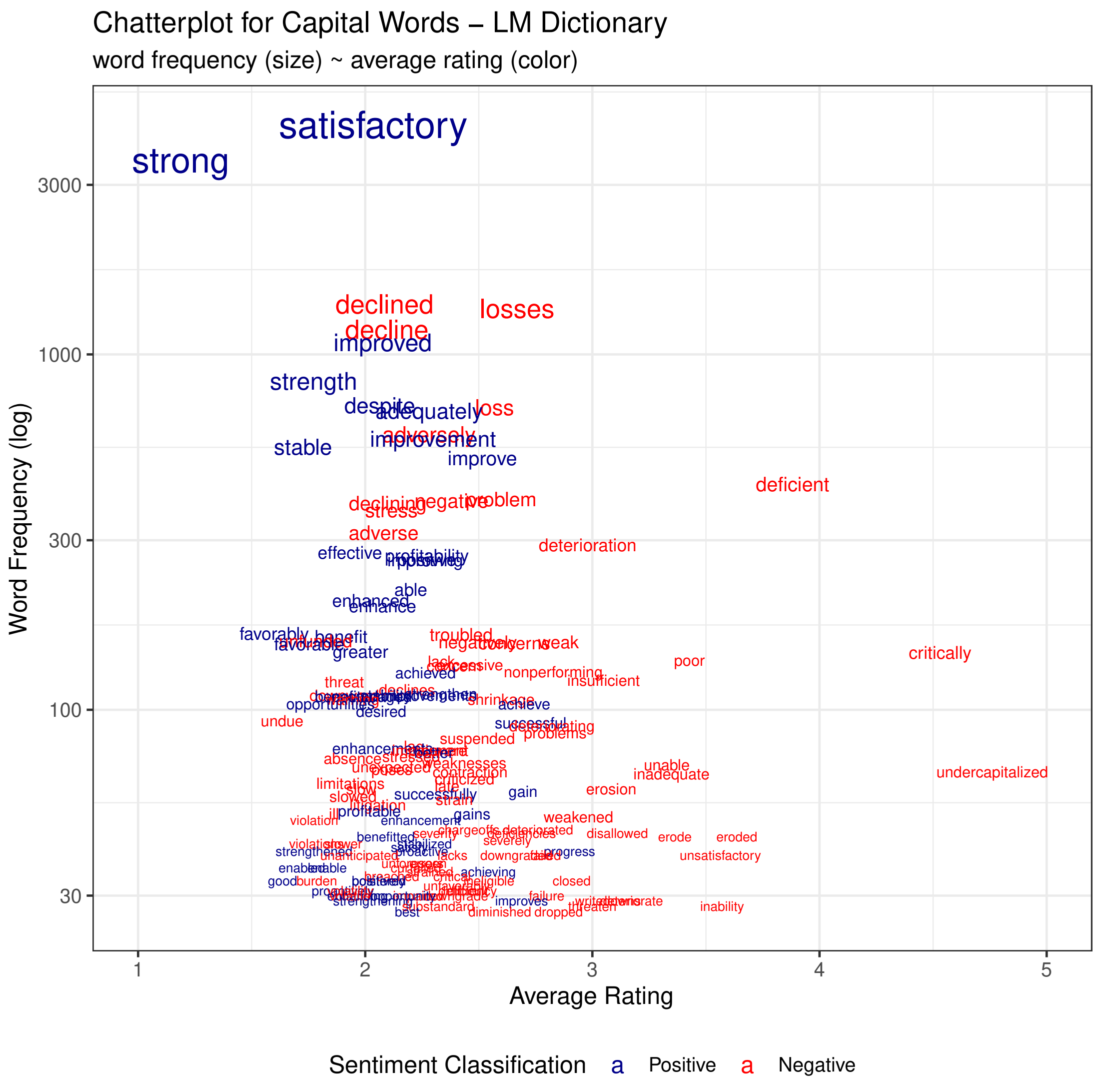
Figure A.5: Frequencies and Average Ratings associated with Words in the LM Dictionary in Bank Exams



Note: Data between 2004:Q1 and 2016:Q2.

Source: Confidential bank exams.

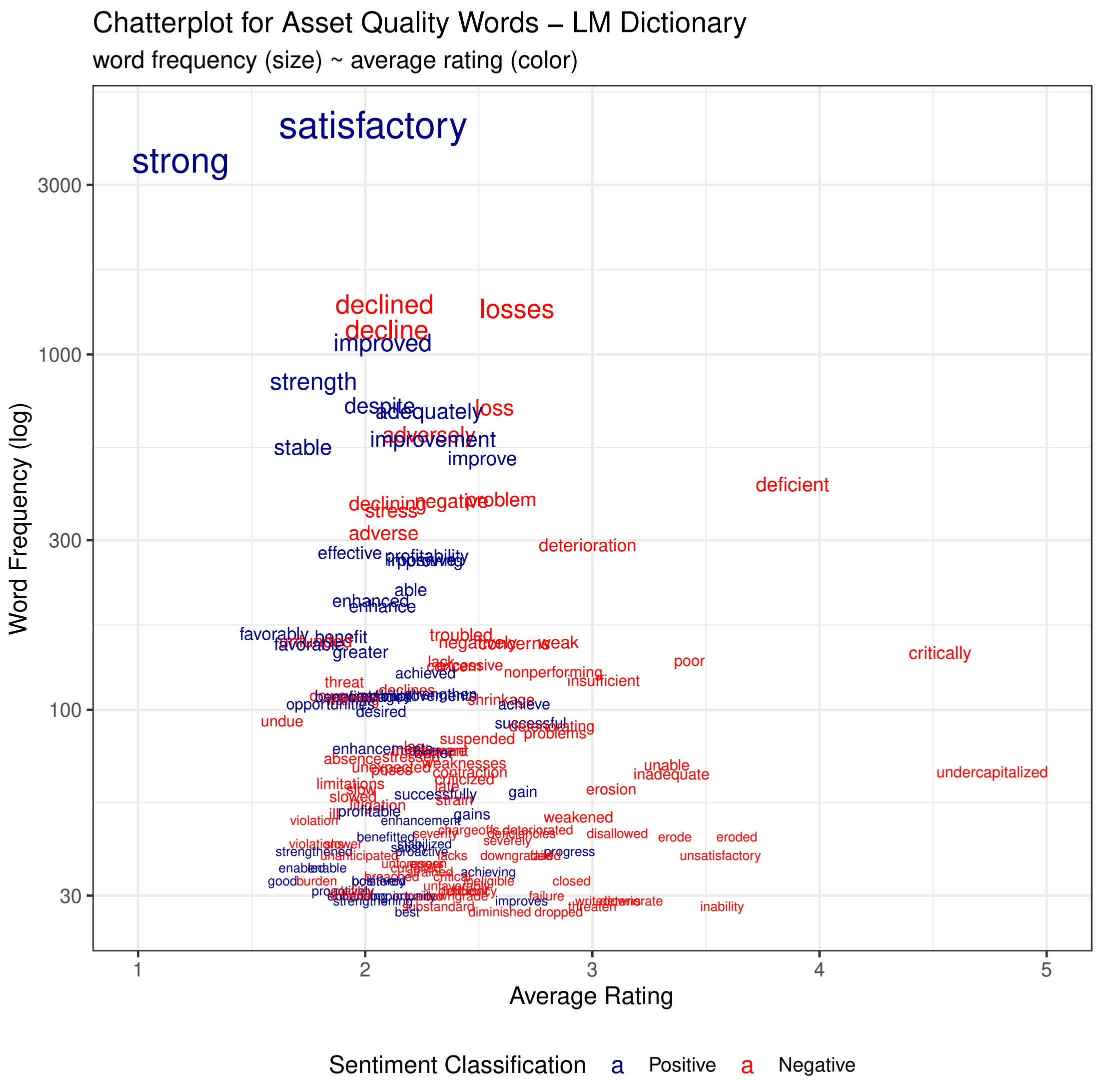
Figure A.6: Frequencies and Average Ratings associated with Words in the LM Dictionary in Capital Sections of Bank Exams



Note: Data between 2004:Q1 and 2016:Q2.

Source: Confidential bank exams.

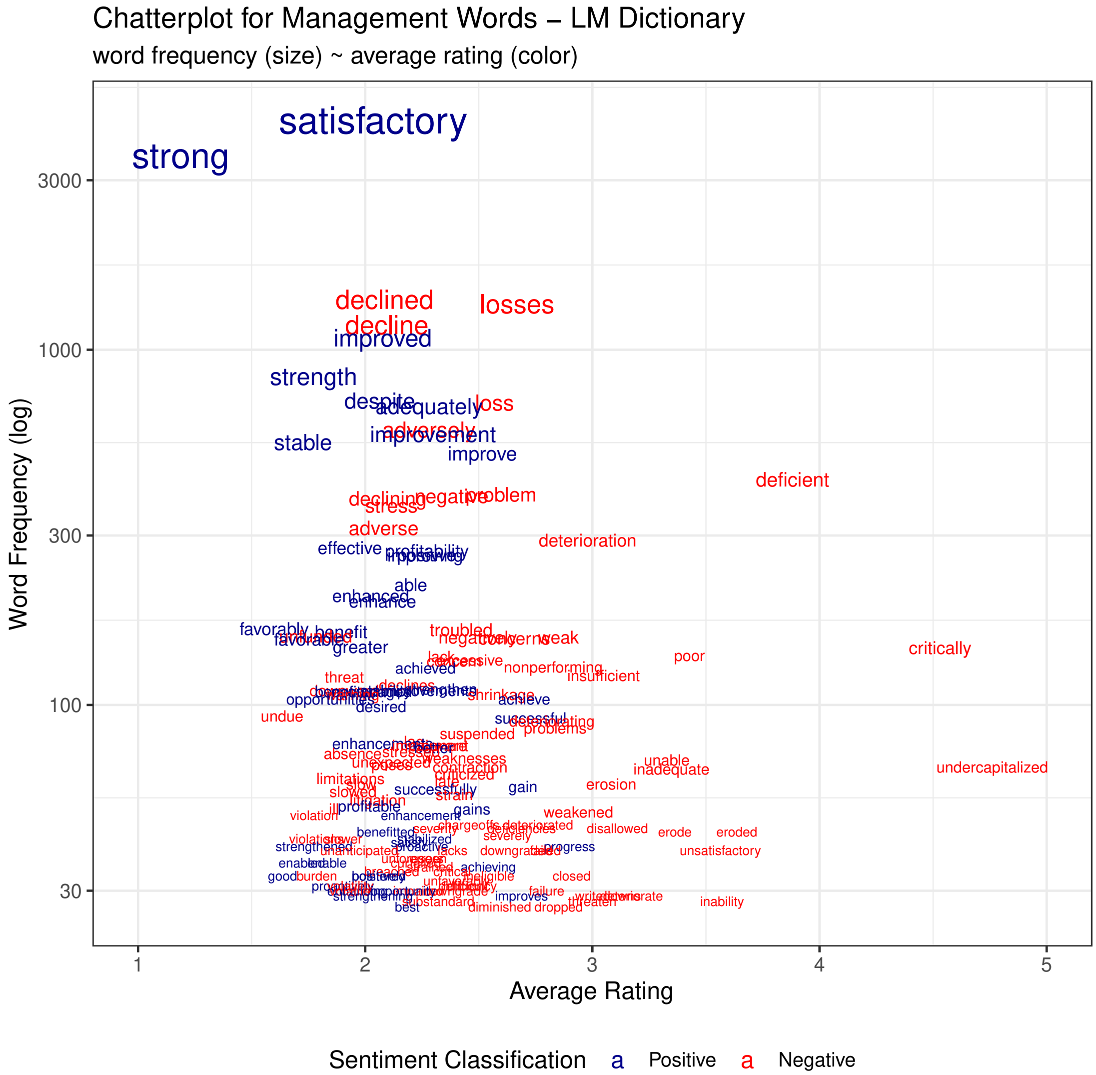
Figure A.7: Frequencies and Average Ratings associated with Words in the LM Dictionary in Asset Quality Sections of Bank Exams



Note: Data between 2004:Q1 and 2016:Q2.

Source: Confidential bank exams.

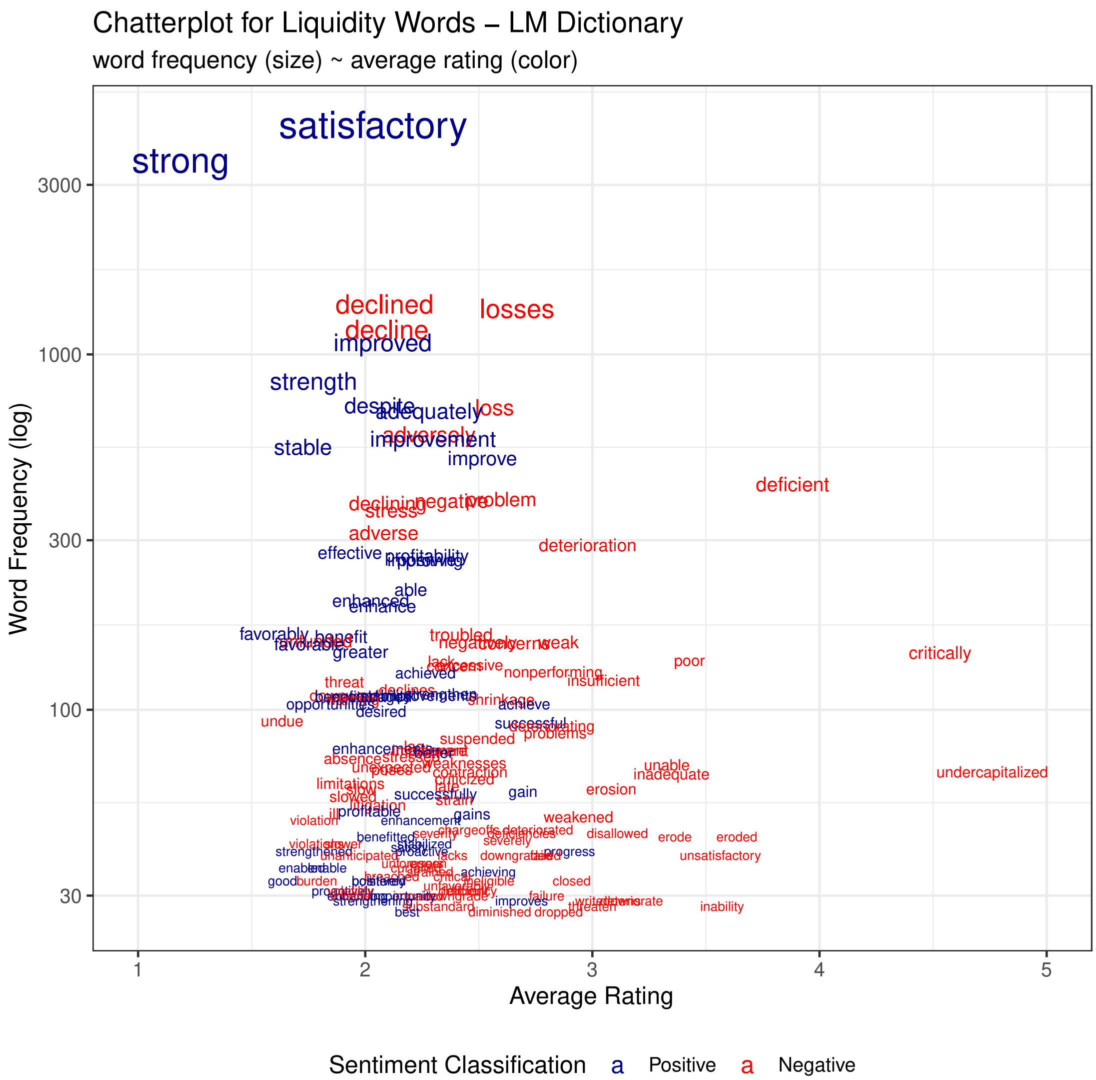
Figure A.8: Frequencies and Average Ratings associated with Words in the LM Dictionary in Management Sections of Bank Exams



Note: Data between 2004:Q1 and 2016:Q2.

Source: Confidential bank exams.

Figure A.9: Frequencies and Average Ratings associated with Words in the LM Dictionary in Liquidity Sections of Bank Exams



Note: Data between 2004:Q1 and 2016:Q2.

Source: Confidential bank exams.

10.2 Community Versus Tiny Banks

The next set of regressions relate bank outcomes and sentiment but separately for banks that have less than \$250 million in consolidated assets and banks that have between \$250 million and \$10 billion in consolidated assets. The sentiment score regression is run separately using the following specifications

$$
\begin{aligned}
\text{outcome}_t &= \rho_{tiny}\ \text{outcome}_{t-1} + \beta_{tiny}\ \text{sentiment}_{i,c,t-1} + \gamma_{tiny}\ log(\text{assets}_{t-1}) \\
&\quad + \Sigma_{n=1}^{4}\ \psi_n\ \text{CAMELS dummy}_{i,c,n,t-1} + \phi_t + \epsilon_{i,t}, \text{if assets} < 250 \text{ million and} \\
\text{outcome}_t &= \rho_{cmty}\ \text{outcome}_{t-1} + \beta_{cmty}\ \text{sentiment}_{i,c,t-1} + \gamma_{cmty}\ log(\text{assets}_{t-1}) \\
&\quad + \Sigma_{n=1}^{4}\ \psi_n\ \text{CAMELS dummy}_{i,c,n,t-1} + \phi_t + \epsilon_{i,t}, \text{if assets} \geq 250 \text{ million and } < 10 \text{ billion.}
\end{aligned}
$$

The CAMELS dummy and the sentiment score are based on the relevant section.

The coefficient of interest is still β. The results are summarized in table A.1.

All of the results are similar between the two size groups. This suggests that the examination process is consistent across banks with \$10 billion in consolidated assets or less.

Table A.1: Regressions–Community and Tiny Separately

This table shows the coefficients and standard errors on the lagged sentiment score. Banks are classified as "community" or "tiny" based on total assets. Tiny banks have assets less than or equal to 250 million dollars. Community banks have assets greater than 250 million dollars and less than 10 billion dollars. The specifications follow this structure:

$$
\begin{aligned}
\text{outcome}_t &= \rho_{tiny}\ \text{outcome}_{t-1} + \beta_{tiny}\ \text{sentiment}_{i,t-1} + \gamma_{tiny}\ log(\text{assets}_{t-1}) \\
&\quad + \Sigma_{n=1}^{4}\ \psi_n\ \text{CAMELS dummy}_{i,n,t-1} + \theta_i + \phi_t + \epsilon_{i,t}, \text{if assets} < 250 \text{ million and} \\
\text{outcome}_t &= \rho_{cmty}\ \text{outcome}_{t-1} + \beta_{cmty}\ \text{sentiment}_{i,t-1} + \gamma_{cmty}\ log(\text{assets}_{t-1}) \\
&\quad + \Sigma_{n=1}^{4}\ \psi_n\ \text{CAMELS dummy}_{i,n,t-1} + \theta_i + \phi_t + \epsilon_{i,t}, \text{if assets} \geq 250 \text{ million and } < 10 \text{ billion.}
\end{aligned}
$$

The CAMELS dummy and the sentiment score are based on the relevant section.

		(1)	(2)	(3)	(4)
		Polar		Valence	
		Community	Tiny	Community	Tiny
Composite					
	MRA/MRIA Sum	-8.739***	-7.570***	-16.52***	-14.09***
		(1.071)	(0.808)	(2.551)	(2.038)
	MRA/MRIA Dummy	-0.681***	-0.710***	-1.335***	-1.365***
		(0.0723)	(0.0617)	(0.157)	(0.143)
Capital					
	Tier 1 Ratio	0.417	0.342	1.346	1.388
		(0.309)	(0.409)	(0.943)	(1.671)
	CET1 Ratio	0.410	0.375	1.365	1.395
		(0.303)	(0.401)	(0.924)	(1.632)
Asset Quality					
	Loan Loss Provisions/Loans	-0.272***	-0.232***	-0.696***	-0.742***
		(0.0551)	(0.0460)	(0.137)	(0.129)
	4-qtr Net Charge-offs/Loans	-0.109***	-0.0755***	-0.288***	-0.244***
		(0.0193)	(0.0188)	(0.0428)	(0.0507)
	Delinquent Loans/Loans	-0.970***	-0.617***	-2.785***	-1.823***
		(0.204)	(0.226)	(0.480)	(0.562)
Management					
	MRA/MRIA Sum	-1.758***	-2.392***	-3.990***	-6.362***
		(0.487)	(0.413)	(1.355)	(1.329)
	MRA/MRIA Dummy	-0.219***	-0.260***	-0.532***	-0.640***
		(0.0348)	(0.0305)	(0.0925)	(0.0902)
Earnings					
	Weighted 4-qtr ROA	0.238***	0.290***	0.636***	0.819***
		(0.0293)	(0.0268)	(0.0818)	(0.0743)
	Weighted 4-qtr PPNR/Assets	0.330***	0.303***	0.871***	0.810***
		(0.0404)	(0.0297)	(0.111)	(0.0781)
Liquidity					
	Securities/Assets	0.593*	0.219	2.731**	0.772
		(0.335)	(0.438)	(1.289)	(1.446)
	(Cash+Securities)/Assets	-0.321	-0.341	-0.927	-2.918*
		(0.428)	(0.499)	(1.540)	(1.667)

Robust standard errors are in parentheses. *** p<0.01, ** p<0.05, * p<0.1