Supervising Failing Banks

Sergio Correia, Stephan Luck, and Emil Verner *

September 26, 2025

Abstract

This paper studies the role of banking supervision in anticipating, monitoring, and disciplining failing banks. We document that supervisors anticipate most bank failures with a high degree of accuracy. Supervisors play an important role in requiring troubled banks to recognize losses, taking enforcement actions, and ultimately closing failing banks. To establish causality, we exploit exogenous variation in supervisory strictness during the Global Financial Crisis. Stricter supervision leads to more loss recognition, reduced dividend payouts, and an increase in the likelihood and speed of closure. Increased strictness entails a trade-off between a lower resolution cost to the FDIC and reduced credit.

JEL: G01, G21, N20, N24, G28, K23, E44

*Correia: Federal Reserve Bank of Richmond, sergio.correia@rich.frb.org; Luck: Federal Reserve Bank of New York, stephan.luck@ny.frb.org; Verner: MIT Sloan School of Management and NBER, everner@mit.edu.
We thank Rosalind Bennett, Mark Carey, Robert Clark, Harry Cooperman, Thomas Eisenbach, João Granja, Yadav Gopalan (discussant), Bev Hirtle, Kathryn Judge, Joe Mason, Matt Plosser, Ben Ranish, Farzad Saidi, Amit Seru, Adi Sunderam, as well as seminar participants at the Federal Reserve Bank of New York, Board of Governors of the Federal Reserve, Foster School of Business at University of Washington, SITE Financial Regulation, Yale Program on Financial Stability, and UMass Amherst for useful comments. We thank Natalia Fischl-Lanzoni and Tiffany Fermin for excellent research assistance. The opinions expressed in this paper do not necessarily reflect those of the Federal Reserve Bank of New York or the Federal Reserve Bank of Richmond.

1 Introduction

Banking supervision seeks to safeguard the health of the financial system by monitoring financial institutions and enforcing regulatory compliance.[1] In the U.S., government agencies and banks spend considerable resources on bank examinations.[2] The effectiveness of banking supervision, however, is hotly debated. Supervision has been subject to a range of critiques, from failing to anticipate bank failures and crises, to delaying the closure of troubled banks, to imposing overly restrictive credit conditions.[3] Understanding the benefits and shortcomings of banking supervision is especially important in light of recent public discussions about possible reform and consolidation of the U.S. supervisory and regulatory landscape.[4]

This paper empirically examines how banking supervision contributes to the financial health of the banking system. Our analysis exploits confidential supervisory data covering all commercial bank examinations in the U.S. between 2000 and 2023. We document that supervisors anticipate most bank failures with a high degree of accuracy at short horizons. Supervisors play a significant role in ensuring that banks' financial statements recognize losses, that troubled banks are subject to enforcement actions, and that banks that are sufficiently undercapitalized and deemed unviable are closed. To establish the causal effect of stricter supervision, we analyze a natural experiment during the 2008 Global Financial Crisis (GFC). We find that banks exogenously subject to stricter supervision during the GFC recognized more losses, increased retained earnings, were *more* likely to

[1] Bank supervision consists of the "monitoring, inspecting, and examining" of banks "to assess their condition and their compliance with relevant laws and regulation." Supervision is notably distinct from bank regulation, which governs "the operations, activities, and acquisitions" of banks (see Board of Governors, 2005).

[2] For instance, Baer (2024) estimates that the OCC employs over 3,000 full-time examiners, while the Federal Reserve employs more than 2,300 examiners. The FDIC has roughly 6,300 total employees as of 2024, though its examiners are not separately disclosed.

[3] For examples of these critiques, see, among others, Financial Crisis Inquiry Commission (2011), Barth et al. (2012), Cole and White (2017), Barr (2023), Bernanke et al. (2019), Zingales (2023), Seru (2025), and Bessent (2025). For an analysis of the behavior of bank supervisors in the run-up to the March 2023 banking turmoil, see Gopalan and Granja (2024).

[4] See, for instance, this recent article in the *Wall Street Journal* from February 11, 2025.

fail, and were closed more quickly. Elevated strictness lowered the cost of failure to the FDIC but also led to reduced credit in the short term.

Taken together, our findings suggest that a key role of banking supervision is to anticipate, monitor, and discipline failing banks. In the presence of deposit insurance, banks are insulated from market discipline (Diamond and Rajan, 2001; Calomiris and Jaremski, 2019; Cucic et al., 2024). This can allow insolvent banks to operate for longer than may be optimal. Our findings suggest that supervision can safeguard the health of the financial system by leading banks to recognize losses faster and, if necessary, closing insolvent banks. Timely intervention can reduce the scope for gambling for resurrection, minimize credit misallocation, and reduce the cost of failures (Kareken and Wallace, 1978; Kane, 1989b; Caballero et al., 2008; Blattner et al., 2023). At the same time, our findings suggest that strict supervision can tighten regulatory capital constraints and reduce credit as a consequence. This finding, in turn, suggests that benefits from stricter supervision such as increased transparency and the closing of insolvent banks need to be carefully traded off against the potential cost of reducing credit.

The first part of the paper provides facts about how banking supervision operates in financially distressed banks. We begin by asking: What do supervisors know about troubled banks? We document that supervisory CAMELS ratings[5] strongly predict failure at short horizons. The high predictability of failure within one year based on CAMELS ratings is in line with supervisors being well informed about a bank's distress ahead of failure and bank closures often directly following a supervisory decision.

Given that supervisory ratings capture the increased risk of bank failures, what actions do supervisors take in failing banks? We document four types of actions. First, supervisors devote more resources to troubled banks. The frequency and length of bank examinations more than double in the run-up to failure. Moreover, failing banks are more likely to face federal supervisors, which Agarwal et al. (2014) show are stricter than state

[5]Supervisory CAMELS ratings assess a bank's financial health based on six criteria: capital, assets, management, earnings, liquidity, and sensitivity to market risk.

supervisors.

Second, supervisors play a crucial role in requiring banks to recognize losses. We document that the probability that a bank revises a previously submitted financial statement shoots up right after bank examinations. This pattern is particularly pronounced in troubled banks. Such revisions can be material. Financial statement revisions following supervisory exams can explain up to 50% of the decline in a bank's capitalization in the five years before failure.

Third, supervisors are more likely to make use of public enforcement actions in failing banks. More than 80% of the banks that failed between 2000 and 2023 were subject to some public enforcement action by federal agencies before they failed, in contrast to only 35% for the banks that did not fail during this period. Moreover, the most drastic form of enforcement action, Prompt Corrective Action (PCA), is almost never used in non-failing banks, while around 25% of failing banks were subject to PCA before failure. Nevertheless, most failing banks are not subject to PCA before failure, possibly due to relatively stringent formal requirements to trigger PCA.

Fourth, failing banks are typically closed through a supervisory decision and bank closures tend to occur in an orderly fashion through a sale to another bank that assumes all of the failing bank's deposits (Granja et al., 2017). Market forces such as sudden deposit outflows very rarely precede supervisory decisions to close banks, and banks are almost always closed on Fridays as a result of a supervisory decision.[6] However, failures are nonetheless costly, and the FDIC's Deposit Insurance Fund (DIF) tends to realize losses. On average, the FDIC loses 23 cents per dollar of failing bank book assets.

The evidence in the first part of the paper is descriptive and does not establish the causal effects of supervisory actions. After all, supervisory actions and bank outcomes are jointly determined. In the second part of the paper, we exploit exogenous variation

[6] These patterns are in stark contrast to the historical, pre-Federal Deposit Insurance Corporation (FDIC) experience of the U.S. banking system, when banks were closed by bank owners or by deposit runs and losses for depositors were large (Correia et al., 2024). They are also in contrast to the more recent failure of Silicon Valley Bank in March 2023.

in supervisory strictness during the GFC to establish causality. Following Agarwal et al. (2014), we define a stricter supervisor as one who gives a lower rating, given the same current and expected financial performance. Our identification strategy exploits that on-site examinations for some banks rotate between being led by federal and state examiners. Importantly, as noted above, federal examiners tend to be stricter supervisors (Agarwal et al., 2014). Some banks on this "rotating schedule" were thus quasi-randomly assigned to a relatively stricter supervisor in the early phase of the GFC. Being assigned a stricter supervisor early in the GFC, in turn, determined the subsequent supervisory strictness throughout the GFC, as supervisors were resource-constrained at the height of the GFC and forced to devote most of their resources on banks already identified as troubled.

We find that being exogenously assigned to stricter supervision has meaningful effects on bank outcomes during the crisis. Otherwise identical banks that exogenously receive stricter supervision are more likely to receive CAMELS downgrades. They also recognize more losses and report lower equity-to-asset ratios. Stricter supervision thus leads to a more conservative assessment of the asset values underlying the calculation of book equity. This translates into more enforcement actions and a substantial reduction in dividend payments for exposed banks. The reduction in payouts implies that stricter supervision likely raises a bank's true capitalization even if reported book equity is lower.

Further, we find that banks subject to stricter supervision are more likely to be closed in the GFC. High supervisory strictness leads to a 3.2 percentage point higher chance of failure during the GFC. This is a considerable effect given an unconditional probability of failure of 3.4 percentage points in our main sample. Thus, our findings suggest that the financial conditions of some banks in the U.S. banking system warranted bank closure under some prevailing supervisory standards, but the banks were allowed to continue to exist. This effect is primarily driven by banks that are more exposed to the collapse in the housing market.

Stricter supervision results in the benefit of a lower costs of failures to the FDIC's

Deposit Insurance Fund (DIF). Banks subject to stricter supervision that end up failing impose around 9% lower losses relative to assets to the FDIC's DIF. This is a sizable effect, as the cost associated with the average failure is about 23% of assets. There are at least four explanations for the lower cost. First, by increasing retained earnings, banks subject to stricter supervision enter failure effectively better capitalized. Second, stricter supervisors, at the margin, likely close relatively healthier banks. Third, banks that are subject to more supervisory scrutiny are closed around one year faster than those that are subject to low supervisory scrutiny, possibly preventing value-reducing gambling for resurrection.[7] Finally, stricter supervision may increase the quality of financial information provided to investors, which may reduce frictions in the FDIC's process of selling failed banks (Granja, 2013).

At the same time, we find that stricter supervision reduces credit growth during the GFC, consistent with tighter regulatory capital constraints. While some lending reduction may reflect reduced evergreening, we observe lower lending across all loan categories, including C&I loans, which had lower default rates during the crisis. Stricter supervision also leads to a reduction in bank employment and the number of branches, indicating a contraction in the inputs to banking. This finding suggests that stricter supervision potentially entails a risk of reduced credit in a crisis (Bernanke et al., 1991).

Taken together, our findings show that supervision is important in monitoring banks, requiring them to realize losses, and shrinking or even closing unviable banking operations following the realization of a negative asset shock. These findings emphasize the importance of supervision for the enforcement of regulatory compliance, loss recognition, and capital adequacy. The causal estimates suggest that making supervision more lenient allows banks to operate longer with undetected losses, potentially allowing nonviable banking operations to continue to exist and gamble for resurrection, with implications

[7] The difference in the timing of bank closures raises the possibility that the reduction in the loss for the DIF stems from differences in the liquidity in the market for failed banks, which may plausibly vary throughout the GFC. We provide additional evidence to show that market liquidity is not the driver of the reduction in losses for the DIF.

for the cost of bank failures to the DIF but also with possible large ramifications for the real economy (see, e.g., Caballero et al., 2008). However, some temporary forbearance can nonetheless be optimal as a tool to support credit during a crisis.

Related Literature. Our paper relates to a rich empirical literature that studies the effects of banking supervision on banking market outcomes (see, e.g., Hirtle and Kovner, 2022, for an overview). The most closely related paper is Agarwal et al. (2014), which studies the differences between federal and state regulatory agencies. Agarwal et al. (2014) show that when comparing federal and state regulator supervisory ratings within the same bank, federal regulators are systematically tougher in their assessment of bank riskiness. We build on their original findings to construct our instrument in Section 4 and use the instrument to document new results about the effect of stricter supervision on troubled banks during the GFC.

Several existing studies use quasi-natural experiments to identify the effects of banking supervision on bank risk-taking and bank lending. Granja and Leuz (2024) exploit the formal closure of the thrift supervisor (OTS) to analyze the effects of supervision on bank lending and bank management. They find that supervision can act as a catalyst for operational changes that correct deficiencies in bank management and lending practices, leading to increased lending. Hirtle et al. (2020) find that banks that receive more supervisory attention hold less risky loan portfolios, are less volatile, and are less sensitive to industry downturns but do not have lower growth or profitability. Kandrac and Schlusche (2020) use variation in supervisory attention to assess the effect of supervision on financial institutions' willingness to take risk and document that supervision can reduce bank risk-taking and lower the chance of bank failures. Altavilla et al. (2020) use credit registry from multiple European countries and show that supervision reduces credit supply to firms with high credit risk but strengthens credit supply to firms without loan delinquencies. Ivanov and Wang (2024) show that banks decrease credit commitments following supervisory downgrades, and Ivanov et al. (2023) show that bank lending rates drop considerably when supervisory scrutiny recedes.

Our paper complements these existing papers by providing causal evidence that banking supervision is key in monitoring and imposing discipline on banks that have become troubled, usually as a consequence of realized credit risk. Thus, our paper focuses less on the role of supervision in affecting bank business decisions outside of financial distress but rather emphasizes that a key role of supervision is to identify distressed institutions and to help avoid "kicking the can down the road" by preventing under-capitalized banks to continue to exist (Kane, 1989a; Kroszner and Strahan, 1996; Acharya et al., 2021). Thereby, our paper also contributes to the work that studies "gambling for resurrection" (see, e.g., Ben-David et al., 2019) and the empirical literature studying the importance of deposit insurance in affecting market discipline (Calomiris and Jaremski, 2019; Cucic et al., 2024).

The specific finding on the role of supervisors in auditing bank call reports also relates to an important set of papers on the relevance of supervisors to ensure the accuracy of banks' financial statements. We complement and reinforce the findings of Costello et al. (2019), who document that being subject to stricter supervisors leads to a higher chance or revising previously submitted call reports. The difference between our analysis of call revision and that of Costello et al. (2019) is that we observe exam dates and lead agencies (as opposed to bank types) and can thus provide more precise estimates of the effects of exams on financial statement revisions. Further, our findings relate to Granja (2013), who studies the effect of disclosure requirements on the resolution costs of failed banks and finds that more comprehensive disclosure requirements improve the efficiency of the sales of failing banks. Granja (2018) exploits cross-state variation during the National Banking Era in disclosure requirements to show that requirements to report financial statements in local newspapers are associated with greater stability and development of commercial banks. Badertscher et al. (2018) document that call reports elicit economically significant stock price and volume reactions when they are publicly released. Bonfim et al. (2023) use administrative data from Portugal to show that bank on-site inspections reduce banks' incentives to refinance zombie firms. Passalacqua et al. (2021) use administrative data from Italy to show that banks are more likely to reclassify loans as non-performing after unexpected bank examinations.

Finally, our CAMELS rating results contribute to the rich literature on the driver and informational content of supervisory ratings. Gopalan and Granja (2024) study the evolution of supervisory ratings during the 2022-23 hiking cycle to explore whether supervisors understood and acted on the potential implications of interest rate risk. Agarwal et al. (2024) show that CAMELS ratings are in part driven by discretion of individual lead examiners. Wheelock and Wilson (2000) examine the determinants of individual bank failures and acquisitions in the U.S. from 1984 to 1993 and find that CAMELS ratings are predictive of bank failure. Berger and Davies (1998) show that changes in regulatory ratings eventually become priced in by the stock market, partially via changes in assessments of the riskiness of a bank's loan portfolio. Cole and Gunther (1998), Hirtle and Lopez (1999), and Berger et al. (2000) study how quickly the informational content or supervisory ratings decay compared to information contained in publicly available Call Reports or stock prices. Oshinsky and Olin (2005) study what determines the final outcomes of banks that receive a CAMELS rating of 4 or 5 at some point in their history.

Roadmap. The paper proceeds as follows. Section 2 describes the data. Section 3 studies the extent to which supervisors anticipate bank failures and discusses the actions supervisors take for both troubled and non-troubled banks. Section 4 provides evidence from a natural experiment in which we trace the impact of exogenous variation in supervisory scrutiny at the onset of the GFC. Section 5 concludes.

2 Data

We build a data set on bank financial statements and revisions, supervisory examinations and ratings, regulatory actions, and bank failures by combining the following data sources.

First, we use examination dates, examination types, and regulatory CAMELS ratings from safety and soundness examinations conducted by bank supervisors and as reported in confidential data provided by the National Information Center (NIC). The Federal Deposit Insurance Corporation Improvement Act of 1991 requires that banks be subject to regular on-site examinations. For state-member banks (SMBs), exams are conducted by either the Federal Reserve System (FRS) or state regulatory agencies. For non-member banks (NMBs), exams are conducted by both the FDIC and state regulatory agencies. For national banks, exams are conducted by the Office of the Comptroller of the Currency (OCC). On-site examinations result in an assessment of a bank's safety and soundness, which is summarized in the CAMELS rating. The composite CAMELS rating is a bank-level score that takes discrete values between 1 and 5, with 1 representing the best possible rating and 5 being the worst rating (see Appendix A.1 for details on the different rating categories). Supervisors assign CAMELS ratings based on their evaluation of six different aspects: capital adequacy, asset quality, management, earnings, liquidity, and sensitivity to market risk. The records from bank examinations are collected in a confidential database available within the FRS.

Second, we use the Federal Financial Institutions Examination Council (FFIEC) Consolidated Reports of Condition and Income ("Call Report"). These publicly available data provide quarterly information on balance sheets and income statements (Forms FFIEC 031, 041, and 051) on a consolidated basis for all commercial banks operating in the U.S. and regulated by the FRS, the FDIC, and the OCC. We also merge in additional information provided by the NIC on bank charters, founding year, branch network, and primary regulator.

Third, we combine publicly available finalized Call Report data with data on historical revisions on specific Call Report line items. These data allow us to study whether regulatory filings are more likely to be revised in response to examinations. We explain these data in more detail in Appendix A.3. Data on historical revisions of Call Report filings are available after 2007.

Fourth, we complement the Call Report data with the FDIC list of failing banks. This list documents all failures of FDIC member banks from 1934 through 2023. We define "failure" as an instance when a bank is closed and either liquidated or its deposit franchise is sold in part or entirely to another bank.[8] Moreover, we use the FDIC Failure Transactions Database to obtain estimates on the cost of failures for the FDIC insurance fund.

Fifth, to measure regulatory actions, we use publicly reported enforcement actions for banks. These include cease-and-desist orders, written agreements, and PCA. Enforcement actions are provided by all three federal banking supervisors—the FDIC, the FRS, and the OCC—on their respective websites. We do not have data on enforcement actions by state bank regulatory authorities.

Finally, to create proxies of bank-level exposure to the GFC, we use information from the FDIC Summary of Deposits, which allows us to track a bank's geographic footprint. We combine this with CoreLogic house price indexes to capture county-level house price growth.

Our main sample is quarterly from 2000 through 2023. In total, we observe more than 200,000 bank examinations, more than 17,000 enforcement actions, and quarterly financial statements for around 10,000 unique banks, of which around 550 fail over the same period. It is important to emphasize that our analysis exploiting exogenous variation in supervisory strictness relies on variation across state-chartered institutions. Thus, for

[8] The FDIC definition of failure also includes "open bank assistance," where the FDIC provides financial assistance to the bank under a systemic risk exception to prevent failure. Our findings are robust to using this broader definition.

this analysis, we drop all OCC-regulated national banks.[9] Moreover, this identification strategy requires banks to qualify for the rotating exam schedule, which also requires us to drop commercial banks with more than $10 billion in assets. Hence, to a large extent, the findings in our paper are about the role of supervision in relatively smaller, state-chartered "community banks" rather than large and complex banking institutions or super-regional banks.

3 Banking Supervision and Failing Banks

In this section, we investigate two questions. First, do supervisory ratings reflect the risk of increased bank failures? Second, what actions do supervisors take in troubled banks?

3.1 Do Supervisory Ratings Reflect the Risk of Bank Failures?

At the conception of CAMELS, a high rating was meant to "include institutions rated as having a high probability of immediate or near term failure" (Federal Reserve Bank of Dallas, 1979). The failure of Silicon Valley Bank (SVB) in March 2023 raised concerns about supervisors' ability to identify bank distress and anticipate potential bank failures. SVB had a CAMELS rating of 2 ("Satisfactory"), leading to the criticism that supervisors were not aware of its risk exposures (Board of Governors, 2023). This raises the question: Do supervisory ratings reflect the increased chance of bank distress before failure? Or are supervisors typically blindsided by failures?

To understand this question, we start by studying the dynamics of CAMELS ratings and basic measures of bank health in failing banks. We estimate variants of the following

[9] Of the around 10,000 unique banks, roughly 25% are national banks, 60% are non-member state banks, and around 15% are FRS-member state banks.

specification for the sample of failing banks:

$$y_{b,t} = \alpha_b + \sum_j \beta_j \times \mathbf{1}[\text{TimeToFail}_{b,t} = j] + \epsilon_{b,t}, \tag{1}$$

where $y_{b,t}$ is a bank outcome such as its CAMELS rating, $\mathbf{1}[\text{TimeToFail}_{b,t} = j]$ is an indicator variable equal to one if a bank fails in j periods from time t, and α_b is a bank fixed effect. The sequence of coefficients $\{\beta_j\}$ measures the evolution of a given outcome in the run-up to failure. The coefficient is normalized to zero ten years before failure.

Figure 1 shows that bank failures are associated with a gradual deterioration of bank solvency in the five years before failure. For the typical failing bank, loan losses translate into declining net income, leading to a reduction in equity-to-assets by about 7 percentage points in the run-up to failure.

Figure 1 further shows that the average CAMELS composite rating of failing banks rises gradually by 2.5 points in the five years before failure.[10] In terms of magnitudes, the increase in the CAMELS rating is considerable. As a benchmark, the unconditional average CAMELS rating is around 1.8 (see Table 1). In our sample of all bank-quarter observations from 2000 through 2023, 32% of observations have a rating of 1 ("strong"), 57% a rating of 2 ("satisfactory"), and 7% a rating of 3 ("less than satisfactory"). Only 3% of all bank-year observations have a CAMELS rating of either 4 ("deficient") or 5 ("critically deficient"). On average, failing banks receive a CAMELS rating of 3 during the five years preceding failure, rising to an average of 4.5 in the last year before failure—with 70% of banks having a rating of 5 and only 7% with either a rating of 1 or 2 (see Table 1). Thus, failing banks are assessed as being much riskier than the average bank, and the rating of failing banks more than doubles, increasing from an average of 1.8 to 4.5 in the

[10] Panel (a) of Figure C.1 in the appendix further shows that the subcomponents of the composite CAMELS ratings increase across the board. In line with a bank's solvency deteriorating due to realized credit risk, the increases are most common in the CAMELS ratings for capital, asset quality, and earnings. Liquidity ratings move up somewhat more slowly. Further, note that the increase in the CAMELS rating is about the same across different size categories, as shown in Panel (b) of Figure C.1.

year leading up to failure.

Table 1: *Distribution of* CAMELS *Ratings across Failing and Surviving Banks.*

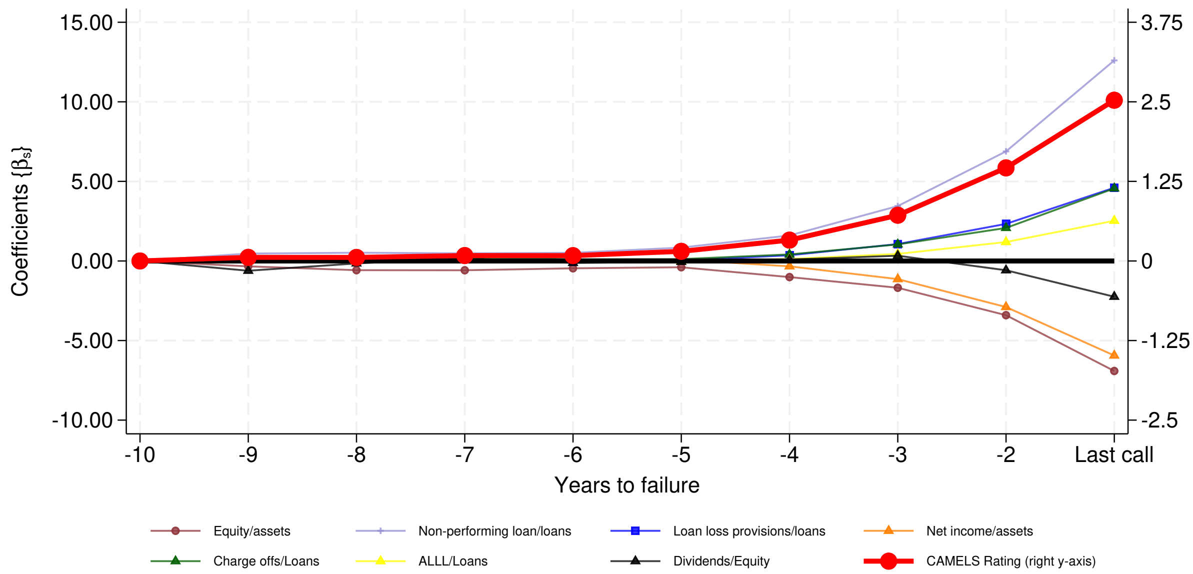
Average CAMELS	Share with CAMELS = X				
	X=1	X=2	X=3	X=4	X=5
A. All banks					
1.83	0.32	0.57	0.07	0.02	0.01
B. Failing banks (within 5 years of failure)					
3.03	0.06	0.46	0.12	0.13	0.24
C. Failing banks (within 1 year of failure)					
4.50	0.01	0.06	0.06	0.17	0.70
D. Surviving banks					
1.81	0.33	0.57	0.07	0.02	0.01

Notes: This table shows the average CAMELS rating and the distribution of CAMELS ratings for our main sample from 2000-2023 (Panel A), failing banks within 5 years of failure (Panel B), failing banks within 1 year of failure (Panel C), banks that have not failed before the end of the sample (Panel D).

To understand whether supervisory CAMELS ratings capture the near term probability of failure, Figure 2 plots the probability of failure over the next three years as a function of the composite CAMELS ratings at time t. The figure shows that the future probability of failure rises strongly in CAMELS ratings. Banks with a CAMELS rating of 1 or 2 have a near-zero probability of failure. Thus, failures like SVB where a bank has a CAMELS of 1 or 2 before failure are rare. Banks with a rating of 4 have a 6 percent probability of failure over the next three years, while the probability shoots up to 30% for banks with a CAMELS of 5. Hence, a bank with a CAMELS rating of 5 is more than 80 times more likely to fail than a bank with a CAMELS rating of 1.

In Appendix D.1 in the Appendix, we further evaluate the predictive power of supervisory ratings for bank failure more formally. We establish that the Area Under

Figure 1: ***Losses, Solvency, and*** **CAMELS** ***Ratings of Failing Banks: 2000-2023***



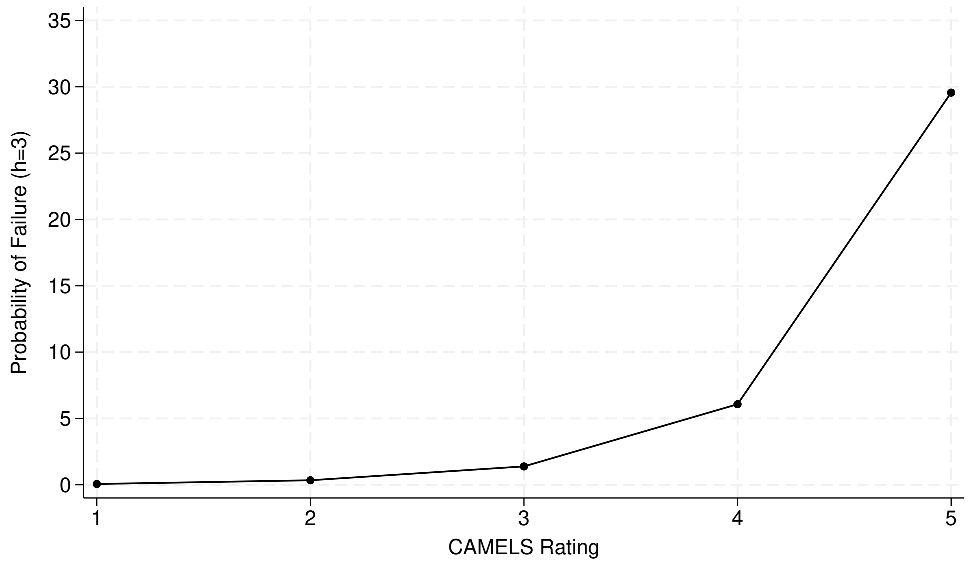
Notes: The figure presents the sequence of coefficients from estimating Equation (1), where the dependent variable is the ratio indicated in the figure legend (left y-axis) or the CAMELS rating (right y-axis). The sample is restricted to failing banks that fail between 2000 and 2023 and to the ten years before they fail.

the Receiver Operating Characteristics Curve (AUC) is 96% at the one-year horizon. A model based on confidential CAMELS ratings can identify around 75% of true positives with less than 1% false positives and close to 90% of true positives with around 5% false positives. We also show that CAMELS ratings tend to be more predictive of failure than financial metrics constructed from Call Report data.[11]

The rise in the CAMELS rating in failing banks in Figure 1 and the high predictive power of CAMELS ratings for failures evidenced in Figure 2 is a clear indication that supervisors typically recognize that failure is highly probable in the period immediately before bank closure. This reflects both the fact that supervisors tend to be well informed about a bank's trouble before failure and the fact that, as discussed in more detail below, bank closures are in part a supervisory decision.

[11]However, we also find that CAMELS ratings underperform simple financial metrics that capture bank performance in predicting bank failures at longer horizons. Thus, the high predictive power of CAMELS for failure is restricted to the near term.

Figure 2: ***CAMELS Ratings and Future Bank Failures***



Notes: This figure plots the probability of bank failure some time between $t+1$ to $t+3$ against a bank's supervisory CAMELS rating at time t.

3.2 What Actions Do Supervisors Take in Failing Banks?

Supervisory ratings anticipate the risk of bank failure, indicating that supervisors are well aware of the increased risk of bank failure. What actions do supervisors take when they realize a bank is in trouble and at risk of failure? We first document that supervisors pay more attention to troubled banks. Second, we show that supervisors take an active role in shaping banks' financial statements, especially for troubled banks. Third, we document enforcement actions. And, finally, we show that supervisors tend to be decisive in determining when a failing bank is closed.

On-site Examinations

We start out by showing that supervisors devote additional attention to troubled or failing banks. As previously noted, banks chartered under state law are supervised by either the Federal Reserve (SMBs) or the FDIC (NMBs), as well as by state regulatory agencies. Federal guidelines require banks to be subject to a scheduled on-site examination every

12 or 18 months.[12] To reduce the regulatory burden on banks under the dual supervision system, federal supervisors can substitute a federal examination with a state examination. Thus, these state-chartered banks can, if certain conditions are fulfilled,[13] be on a "rotating" schedule under which a federal exam is followed by a state exam and vice versa. Figure 3 plots the time elapsed between two consecutive exams in the sample of banks eligible for the rotating schedule. The solid green line confirms that banks eligible for the rotating schedule typically undergo an exam every 12 or 18 months. The dashed lines in Figure 3 show that failing banks, in contrast, are subject to more frequent exams than banks on the rotating schedule. The distribution of days elapsed between exams shifts to the left for banks that are approaching failure.

Table 2 confirms the pattern from Figure 3 and shows that the average on-site examination starts around 415 days after the previous examination. For banks on the rotating schedule, the average days between exams is 470. By contrast, exams in failing banks that are within the last five years of their existence take place 256 days after the most recent previous exam; if a bank is within the last year of its existence, exams tend to take place every 177 days. Thus, while the average bank is subject to less than one exam per year, failing banks can expect around two exams per year right before they fail.

Moreover, for failing banks supervised by both federal and state agencies, federal agencies are more likely to become the lead examiner in the years leading up to failure. Table 2 shows that, across all banks, state supervisors lead around 35% of bank exams.

[12] Federal bank supervisors are required to conduct on-site examinations every 12 months unless their assets fall below a minimum threshold (which has evolved over time and, as of 2023, is at $3 billion), in which case the exams are conducted every 18 months. Further, larger banks ($10 billion or more in assets) are subject to regular joint examinations or continuous examinations. National banks supervised by the OCC are on a different schedule, with exams typically taking place at a higher frequency, often twice within the same year.

[13] According to the FDIC examination manual, "Examinations may be conducted in alternate 12- or 18-month periods if the FDIC determines that a full-scope, onsite examination completed by the appropriate state supervisory authority during the interim period is acceptable. However, such alternate examinations should be accepted only for the following institutions: composite 1- or 2-rated institutions, and stable and improving composite 3-rated institutions if the composite rating is confirmed by an offsite review and no adverse trends are noted from other available information. The length of time between the end of one examination and the start of the next (whether one or both of the examinations are conducted by a state supervisory agency or the FDIC) should not exceed 12- or 18-months" (FDIC, 2022, p. 1.1-5).

However, state supervisors only lead about 14% of the exams for banks within a year of failure, with the remaining 86% led by a federal agency. This evidence suggests that, as banks become troubled, they stop being on the rotating schedule and hence become subject to additional examinations and supervisory scrutiny from the stricter federal supervisors (Agarwal et al., 2014).

Finally, failing banks undergo longer exams. Table 2 shows that the average length of an exam increases by three weeks in the year before failure. At the same time, the share of unscheduled exams increases from more than 5% in all banks to around 14% in such failing banks. Moreover, routine exams become less likely, and exams are more likely to be classified as "Advisory/Visitation."

Taken together, the evidence shows that failing banks receive considerable additional scrutiny from supervisors, in line with supervisors making active decisions on how to allocate their time (Eisenbach et al., 2022). Supervisors conduct more on-site examinations, and troubled banks are more commonly subject to exams led by the relatively stricter federal agencies.

Shaping Public Financial Statements

We next document that supervisors take an active role in shaping the public financial statements of banks through on-site examinations.

Banks have considerable discretion when providing their financial information (Dahl et al., 1998). For instance, banks have discretion when provisioning for potential loan losses or when deciding to charge off loans that have already been provisioned for. This type of discretion, in turn, increases the scope for supervision to ensure the accuracy of banks' financial statements.

One way to understand both the role of banks' discretion in their financial reporting and the role of bank supervision that results from this discretion is to study the revisions of previously submitted financial statements (see, e.g., Costello et al., 2019; Passalacqua

Table 2: *Exam Frequencies, Lead Examiners, and Types of Exams.*

Days between exams		Examiner		Type of exam			Length of exam (days)	
Median	Average	FDIC/FRS	State	Routine	Advisory/Visitation	Other	Median	Average
A. All banks								
448	415	0.60	0.40	0.94	0.05	0.01	24	31
B. Failing banks (within 5 years of failure)								
210	256	0.75	0.25	0.90	0.09	0.01	35	44
C. Failing banks (within 1 year of failure)								
145	177	0.87	0.13	0.85	0.14	0.01	43	51
D. Surviving banks								
462	420	0.59	0.41	0.95	0.05	0.01	24	31
E. Bank eligible for rotating schedule								
518	470	0.54	0.46	0.98	0.01	0.01	23	29

Notes: This table shows the average number of days between supervisory exams, the share of exams led by a federal or state examiner, the type of exam, and the length of the exam (in days). We consider all banks in our main sample from 2000-2023 (Panel A), failing banks within 5 years of failure (Panel B), failing banks within 1 year of failure (Panel C), banks that did not fail by the end of the sample (Panel D), and banks eligible for the rotating schedule (Panel E).

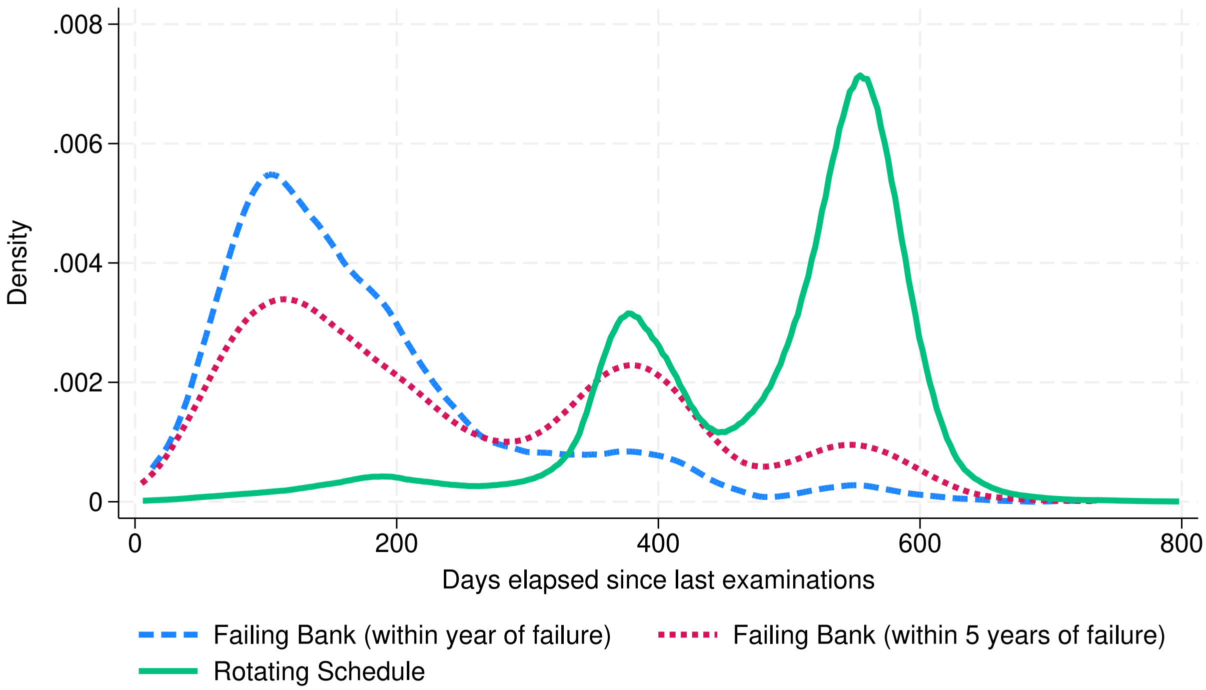
et al., 2021). Banks commonly revise their originally submitted Call Reports. For example, from 2007 through 2023, the probability of a revision to at least one line item is almost 30% (see Table A.1). Importantly, revisions can affect crucial line items such as assets, net income, equity, or loan loss provisions. For instance, net income is revised in 4.6% of all cases and equity in around 5% of all cases.

To establish a relation between revisions of Call Reports and actions of banking supervisors, we test whether banks are more likely to revise their original Call Report submission after on-site examinations. We estimate a model of the following form:

$$\mathbf{1}[\text{Revision}_{b,t}] = \alpha_b + \tau_t + \sum_{j=-6}^{6} \beta_j \times \text{Exam}_{b,t+j} + \epsilon_{b,t}, \tag{2}$$

where $\mathbf{1}[\text{Revision}_{b,t}]$ is an indicator for whether bank b revises either its net income or book equity (from a Call Report submitted before t) in month t, $\text{Exam}_{b,t+j}$ is an indicator

Figure 3: ***Examination Frequency, 2000-2023***

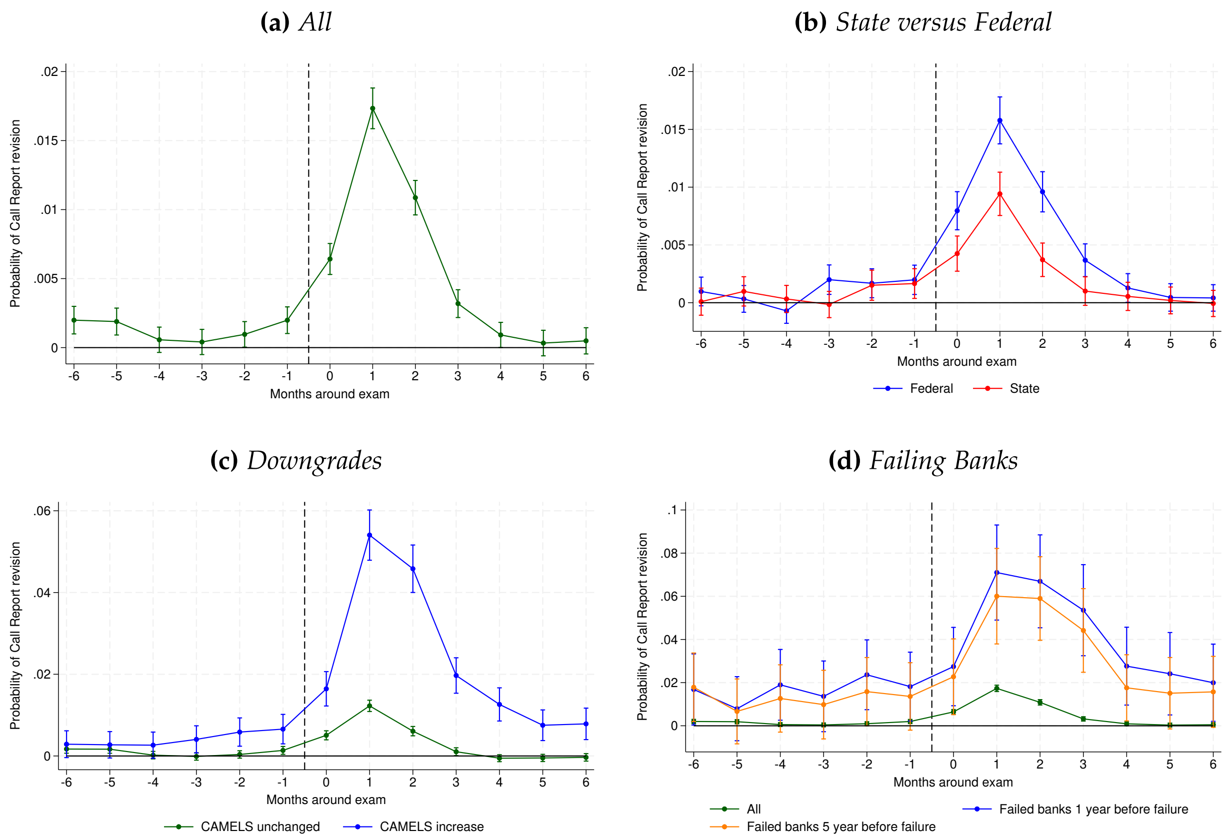


Notes: This figure plots the distribution of the number of days elapsed since a bank was subject to its last commercial bank examination. We restrict the sample to SMBs and NMBs. The green line shows the distribution of exams for banks that qualify for the rotating schedule, the red line shows the distribution of exams for banks that will fail within five years, and the blue line shows the distribution of exams of banks that will fail within one year.

for whether a bank was subject to an examination that started in month $t+j$, α_b is a set of bank fixed effects, and τ_t is a set of calendar-month fixed effects.

Figure 4 plots the set of coefficients $\{\beta_j\}$ from estimating Equation (2) for the six months around an on-site examination. We find that the probability of a revision of a previously submitted Call Report increases slightly in the two months before an examination and then shoots up in the month of and right after an examination (panel (a)). This pattern suggests that banks, in part, anticipate the scrutiny from examiners and exams (recall that exams are announced far ahead), explaining the slight increase in the revision probability before the exam. However, the examination itself is associated with a much higher propensity to revise original submissions. Around six months after the exam, the probability of a revision reverts back to its pre-examination mean.

Figure 4: ***Revisions of Net Income and Equity Reported in the Call Reports around Supervisory Examinations***



Notes: The figure presents the sequence of coefficients from estimating Equation (1), in which the dependent variable is a dummy that indicates whether a bank revises its previously submitted call report line items for net income or equity, or both. The specification includes a set of bank fixed effects. The sample is restricted to all commercial banks and to 2007-2023.

An important challenge in interpreting the patterns in panel (a) of Figure 4 results from the fact that it is unclear whether one can attribute the revision of a previously submitted financial statement to actions taken by supervisors. Revisions may be coincidental to exams rather than caused by them. For instance, banks could be planning on revising their Call Report filing independent of the on-site examination but holding back with resubmissions until the weeks right after the exam, when all types of paperwork need to be processed.

To address this type of identification concern, we consider differences in the revision probability for banks that are on the rotating schedule across exams led by a federal or state agency, similar to the identification in Costello et al. (2019). Given that the exams on the rotating schedule are fully anticipated, any differences in the revision probability following state-led versus federal-led examinations would allow for a causal interpretation in which actions taken by the supervisors are causing banks to revise their financial statements. Indeed, Figure 4 shows that the effect is stronger for exams led by federal examiners compared to state examiners when restricting the sample to banks that are on the rotating schedule. Thus, at least some revisions can be attributed to actions by supervisors.[14] The evidence using the exogenous variation in supervisory strictness during the GFC from the rotating schedule in the next section further reinforces this point.

We also find that revisions following examinations are more likely in troubled banks, banks experiencing a CAMELS rating downgrade, and banks that end up failing (see panels (c) and (d) in Figure 4). Thus, supervisors play an especially important role in the Call Report revisions for banks in financial distress (Gunther and Moore, 2002).

Thus far, we have focused only on whether line items are revised but not on the direction—i.e., the sign—of the revision. Table 3 shows the probability of upward and downward Call Report revisions in any quarter for which a bank examination has taken place in the previous quarter. We distinguish between banks that end up failing during our sample period and those that do not fail. Revisions are considerably more likely in failing banks than in non-failing banks. Further, Table 3 reveals that when failing banks submit revisions, they are more likely to revise upward loan loss provisions, ALLL, and charge-offs. In contrast, they are more likely to revise downward book equity and net income. Remarkably, a bank within a year of failure has a 30% chance of submitting a downward revision to equity and a 27% chance for net income, compared to a 2% and 7% chance of submitting an upward revision of those items, respectively. Likewise, failing

[14]The evidence in Figure 4 panel (b) does not distinguish between whether supervisors are only preventing banks from delaying their statements (and thus the revisions were unavoidable eventually) or whether supervisory actions are causing revisions that would otherwise never happen.

Table 3: ***Probability of Revising Call Report in the Quarter Following an On-Site Examination***

Call Report line items											
ALLL		Provisions		Charge-off		Net Income		Equity		NPL	
Direction of revision											
Up	Down	Up	Down	Up	Down	Up	Down	Up	Down	Up	Down
A. All banks											
0.04	0.01	0.04	0.01	0.02	0.01	0.02	0.06	0.02	0.07	0.02	0.02
B. Failing banks (within 5 years of failure)											
0.19	0.05	0.19	0.03	0.12	0.03	0.05	0.25	0.03	0.28	0.07	0.08
C. Failing banks (within 1 year of failure)											
0.21	0.05	0.21	0.03	0.14	0.03	0.06	0.28	0.03	0.30	0.09	0.11
D. Surviving banks											
0.01	0.01	0.01	0.00	0.01	0.00	0.01	0.03	0.01	0.03	0.01	0.01

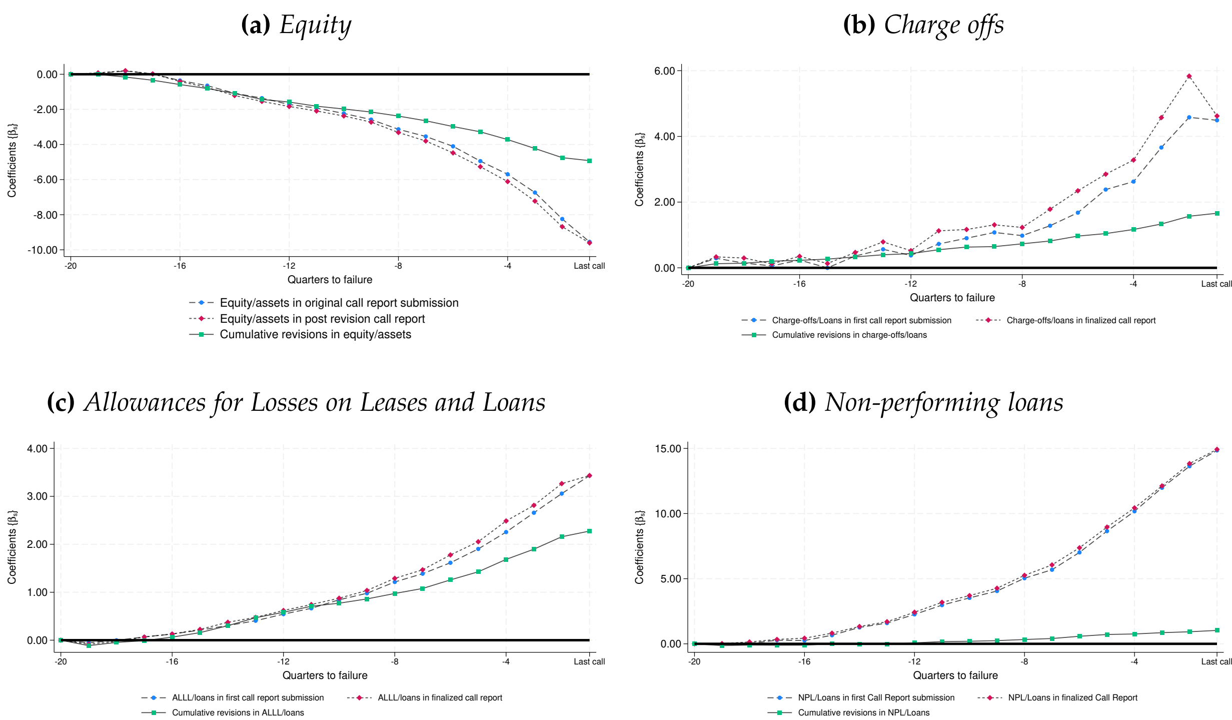
Notes: This table shows the probability of a call report revision for the entire sample of all banks from 2000-2023 (Panel A), for failing banks within 5 years of failure (Panel B), failing banks within 1 year of failure (Panel C) and banks that have not failed before the end of the sample (Panel D).

banks revise their loan loss provisions upward in 18% of cases, compared to a downward revision in only 3% of cases. This finding reinforces the idea that increases in provisions are a key driver of the reduction in reported bank income. Notably, the same is not true for nonperforming loans, which are more likely to decrease after a revision, possibly reflecting supervisors forcing banks to charge off nonperforming loans as well as the lack of discretion in classifying non-performing loans.

The revisions of previously submitted financial statements can be material in troubled banks. Figure 5 plots the dynamics of revisions in failing banks in the five years before they fail. Panel (a) shows that the equity-to-assets ratio tends to fall by around 9 percentage points in failing banks in the five years before they fail. Around 4 percentage points of the decline come after revisions of original Call Report submissions.

Taken together, the findings suggest that banks have considerable discretion in their provisioning practices, which gives supervisors a crucial role in auditing banks and

Figure 5: ***Call Report Revisions in Failing Banks***



Notes: The figure presents the sequence of coefficients from estimating

$$y_{b,t} = \alpha_b + \sum_j \beta_j \times \mathbf{1}[\text{TimeToFail}_{b,t} = j] + \epsilon_{b,t},$$

where $y_{b,t}$ is a bank-level outcome such as the bank's equity-to-assets ratio, $\mathbf{1}[\text{TimeToFail}_{b,t} = j]$ is an indicator that equals one for a bank that will fail in j periods, and α_b is a bank fixed effect. The sample is restricted to failing banks, to the five years before they fail, and to banks that fail after 2007.

ensuring the accuracy of their financial statements.

Public Enforcement Actions

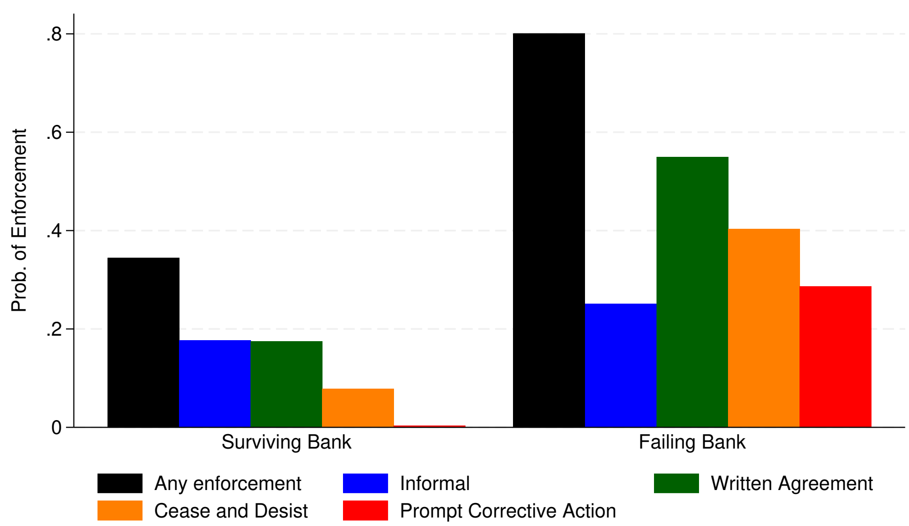
An important supervisory tool is to take enforcement actions. Supervisors can take formal, publicly posted enforcement actions if a bank is in violation of laws, rules, and regulations, or engaging in "unsafe and unsound" practices.[15] Formal enforcement actions range from written agreements and cease-and-desist orders to more drastic measures such as

[15]Supervisors may also undertake informal forms of action by issuing Matters Requiring (Immediate) Attention (see Eisenbach et al., 2017). These actions are confidential and thus do not affect a bank's business model in a way public enforcement action would.

PCA (see, e.g., Curry et al., 1999).[16] As noted above, formal enforcement actions taken by federal supervisors are made public once they become effective.

We next study whether supervisors are more likely to take enforcement actions in troubled banks. Figure 6 shows the probability of being subject to public enforcement action across failing and surviving banks in the 2000-2023 sample. Among all banks that end up failing at some point over that period, at least 80% were subject to some form of public enforcement action before they failed.[17] This probability is considerably lower for banks that have not failed during the same time, of which only around 35% were subject to some form of enforcement action. Enforcement action is more than twice as likely in banks that end up failing. Hence, supervisors act on some of the information they have about troubled banks.

Figure 6: ***Enforcement Actions for Failing and Surviving Banks***



Notes:This figure plots the probability of a bank being at least once subject to public enforcement action between 2000 and 2023, distinguishing between banks that end up failing over the same time period and those that don't fail.

The difference across failing and surviving banks is even more striking when considering the most drastic forms of enforcement. Around 25% of all banks that failed between

[16]PCA requires a bank to issue new equity. It also restricts activities banks can undertake more broadly, including restrictions on interest rates banks can offer on deposits.

[17]Note that we do not observe formal enforcement actions taken by state regulatory agencies, which makes our estimates of the probability of public enforcement action a lower bound.

2000 and 2023 were subject to PCA before they failed, whereas almost no surviving bank was subject to PCA. Notably, even though PCA is considerably more likely in failing banks, the majority of failing banks—even if subject to some formal enforcement action—are not subject to PCA before they fail.

What determines whether a bank is subject to enforcement action? What drives the absence of drastic enforcement action before failure in most failing banks? These questions are especially important in light of the fact that supervisors can anticipate when banks are in trouble, but yet they seem not to make use of the most drastic form of enforcement action. The lack of PCA is also noteworthy, as it is an important pillar of the FDIC Improvement Act of 1991, passed in the wake of the S&L crisis to reduce the cost of bank failures. In Appendix D.2, we document and discuss that a key determinant of enforcement action is whether a bank is undercapitalized in a legal sense.[18]

The fact that enforcement action is closely tied to a bank's capitalization, in turn, has two important implications. First, it offers a potential explanation for the lack of drastic enforcement action, such as PCA. These enforcement actions require banks to be "significantly" or "critically" undercapitalized in a legal sense. However, even as banks' solvency deteriorates for many years in the run-up to failure, they often only become "critically" undercapitalized only right before failure (see Figure D.9). Thus, there may be insufficient time between when a bank becomes undercapitalized and when the bank ultimately fails for supervisors to undertake drastic enforcement action. Book equity measures may also overestimate capitalization in failing banks when banks do not adequately account for losses, especially given the large asset losses in failure, discussed below (Cole and White, 2017).

Second, the close link between a bank's capitalization and enforcement action increases the importance of the auditing function of supervision. As discussed above, a substantial

[18]In principle, supervisors may also hold back on taking enforcement actions even in banks they have identified as problematic because of political or economic constraints (Brown and Dinç, 2005) or distorted incentives (see, e.g., Lucca et al., 2014).

decline in a failing bank's capitalization follows restatements of previously submitted call reports. As we discuss in more detail in Appendix D.2, supervisors influence a bank's financial statement so that enforcement action becomes legally permissible. Without this auditing role, drastic enforcement actions would be even less common.

Bank Resolution

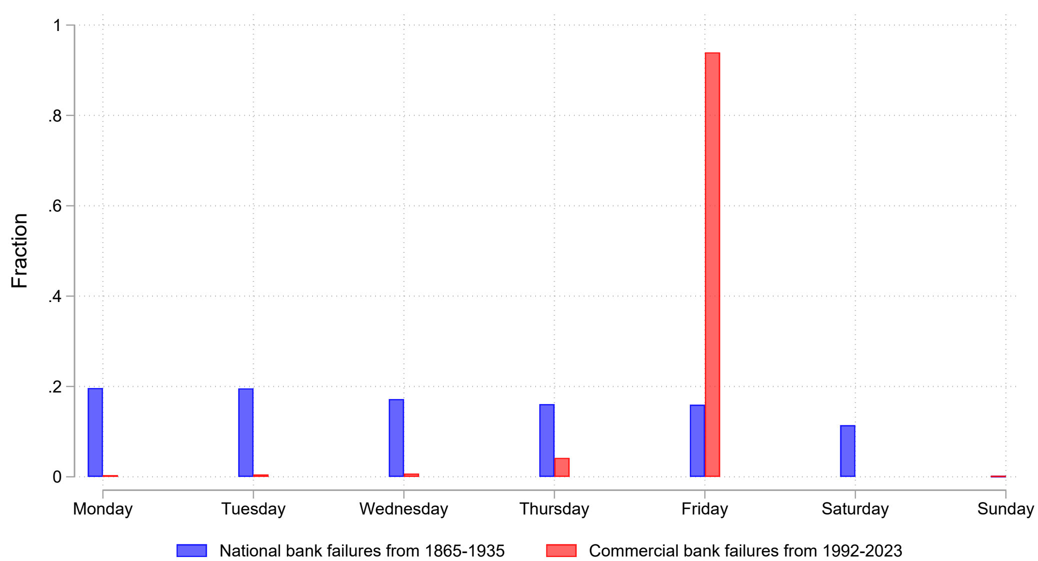
Finally, we provide evidence on the role of bank supervisors in bank resolution. In line with the fact that supervisors are typically not surprised by bank failure, we document that the timing of bank closures typically involves a supervisory decision. To illustrate this point, Figure 7 shows the distribution of failure dates across weekdays. Banks seldom fail on days other than Friday, the FDIC's preferred bank closure day (see, e.g., Walter, 2004). Of all 557 commercial bank failures since 2000, only 28 banks failed on a day different than Friday. Moreover, we find that banks that failed Fridays had on average a 4.9 CAMELS rating before failure. In contrast, those banks that failed on other days of the week only had an average CAMELS rating of 4.4. These patterns indicate that supervisory action almost always determines the timing of a bank closure. Supervisors are typically not blindsided by failure but rather make an active decision to close a failing bank.

In the presence of deposit insurance, market forces such as deposit withdrawals rarely trigger bank closure, consistent with existing empirical evidence that suggests that deposit insurance and other government interventions heavily affect depositor behavior (see, e.g., Martin et al., 2023; Iyer et al., 2019; Cucic et al., 2024).[19] However, this pattern of supervisory decisions governing the timing of bank closures is in contrast to pre-FDIC failures. Using data on bank failures between 1865 and 1935—which includes both the pre-FRS and pre-FDIC eras—provided by Correia et al. (2024), Figure 7 shows that failure

[19]However, we find that, while failures with runs are uncommon in the modern banking system, for the few failures since 1993 in which a failing bank was subject to a deposit outflow exceeding 20%, failure was around 30 percentage points less likely to occur on Fridays.

dates were effectively uniformly distributed across weekdays. This finding indicates that, before federal deposit insurance, bank owners or depositors typically determined the timing of bank closure, while in the contemporary U.S. banking system bank closures are usually a supervisory decision.[20] This comparison highlights how supervisory actions may substitute for the lack of depositor discipline in the absence of deposit insurance (Diamond and Rajan, 2001).

Figure 7: ***Weekday of Bank Failures Pre- and Post-FDIC***



Notes: This figure plots the fraction of bank failures by weekday for both a pre-FDIC sample consisting of failed national banks from 1865 through 1935 and a sample of failed FDIC member banks from 1993 through 2023. Data on national bank failures are from Correia et al. (2024).

Furthermore, when banks are closed, they are almost always purchased by another bank. For instance, more than 500 of the 557 bank failures since 2000 resulted in the purchase of the bank failing by another bank (see Figure C.3 in the appendix). The purchasing bank, in turn, typically assumed all deposits, including uninsured deposits. Failures in which only insured deposits were assumed, or in which the deposit insurance fund required a payout, account for less than 10% of failures since 2000.

The fact that most failing banks' entire deposit franchise (rather than just the insured

[20] Before deposit insurance, insolvent banks were often closed by depositor runs, though Correia et al. (2024) present evidence that runs rarely caused the failure of solvent banks.

deposits) is purchased by another bank implies, in turn, that loss rates on uninsured deposits have been low in the recent history of the U.S., as seen in Table B.6 in the Appendix. For instance, since 2008, the unconditional loss rate on uninsured deposits has been, on average, 3 cents on the dollar. Again, this finding contrasts sharply with the pre-FDIC era, for which Correia et al. (2024) estimate depositor loss rates exceeding 30 cents on the dollar.

These findings, however, do not imply that bank closures are not costly. The FDIC itself tends to realize losses during the resolution, see Table 4. On average, the FDIC DIF loses 23 cents per dollar of assets held at suspension. The FDIC's DIF realizes losses for different reasons. First, losses can result from low recovery rates on the failed bank's investments (James, 1991). Second, losses can arise through inefficiencies when the FDIC sells banks to bidders that are capital or liquidity constrained or not willing or able to pay the market value of the failing bank (Granja, 2013; Granja et al., 2017; Allen et al., 2023a,b). Third, the resolution process itself is costly and requires legal fees.

Table 4: *Cost of Failures to the FDIC Insurance Fund as a Share of Banks Assets at Resolution.*

Era	Average	Median	Share of cost of failure falling within...						N
			≤ 0	$[0, 0.1]$	$[0.1, 0.2]$	$[0.2, 0.3]$	$[0.3, 0.6]$	> 0.6	
1993-2007	0.22	0.15	0.11	0.36	0.28	0.16	0.13	0.07	304
2008-2023	0.23	0.22	0.02	0.15	0.27	0.29	0.27	0.02	542
All	0.23	0.20	0.05	0.23	0.27	0.24	0.22	0.04	846

Notes: This table reports the ratio of the total cost borne by the insurance agency for the failed institution and the failed institutions assets reported at failure. Before 2007, losses are losses to the Bank Insurance Fund (BIF). After 2007, reported losses are to the Deposit Insurance Fund (DIF). For terminated receiverships or assistance agreements, this is total actual losses. For active receiverships or assistance agreements, this is the sum of actual losses to date plus projected losses. The data are from the FDIC's Financial Transaction Database.

4 Causal Effect of Stricter Bank Supervision

In this section, we ask: How does stricter banking supervision causally affect bank-level outcomes during a banking crisis when many banks are exposed to large asset losses? Does stricter supervision lead to faster recognition of losses, more enforcement actions, and potentially more bank closures? Does this stricter supervision come with the benefit of lower cost of failures? Are there downsides to stricter supervision in a crisis, for example, in the form of reduced credit? This section exploits exogenous variation in supervisory strictness during the 2008 Global Financial Crisis and finds that the answer to each of these questions is "yes."

4.1 Identification Strategy

The evidence presented in the previous section is mostly muted on the causal effects of supervisory actions on bank outcomes. Empirically, the effect of supervisory action on bank outcomes is difficult to assess by solely studying descriptive facts. Supervisory scrutiny and supervisory actions are endogenous and typically a response to a bank's financial conditions. For instance, as previously discussed, troubled banks are subject to more frequent examinations and more likely to be subject to an on-site exam led by relatively tougher federal supervisors.

To address this identification challenge, we exploit exogenous variation in exposure to stricter supervision at the onset of the GFC. The GFC provides an appealing setting to study the consequences of stricter supervision. The large downturn in the housing market resulted in a substantial increase in real estate-related losses for banks. Over 500 FDIC-insured banks failed during the crisis. During the crisis, supervisors faced the challenge of identifying losses in the system. Moreover, they faced the trade-off of whether to force troubled banks to promptly recognize losses or engage in forbearance (e.g. Kane, 1989b). Our hypothesis is that stricter supervision in such a crisis is more likely to impose discipline on troubled banks by forcing them to recognize losses and closing insolvent banks. We then study the implications of stricter supervision for the likelihood of failures, cost of failures, and banking activity.

Our identification leverages differences across supervisors from federal and state agencies. State agency-led on-site exams tend to be more lenient than those by federal supervisors, as documented by Agarwal et al. (2014). We make use of the institutional fact that there is quasi-random assignment of examiners via the pre-determined rotating schedule of supervisors at the onset of the GFC. Banks that were on the rotating schedule before the GFC and had a state-led exam before the start of the GFC were more likely to be subject to a federal-led exam early in the GFC for reasons unrelated to their underlying financial conditions.

We estimate bank-level regression models of the following form:

$$y_b = \beta \times \text{High supervisory strictness}_b + \gamma \times X_b + \epsilon_b, \tag{3}$$

where y_b is a bank-level outcome realized during or in the aftermath of the GFC. High supervisory strictness$_b$ is a dummy that takes the value one if bank b is subject to at least one exam led by federal supervisors in the "early" phase of the GFC. We define "early" in the GFC as the 18 months after the failure of Lehman Brothers.[21] Finally, X_b is a set of bank-level control variables.[22]

As noted above, whether a bank is subject to an exam led by a federal supervisor is endogenous and federal supervisors are more likely to lead exams in troubled banks. Therefore, we instrument the "High supervisory strictness" dummy with being on the rotating schedule and the last exam before the GFC being state-led. The identifying

[21] We choose an 18-month window because most banks are on an 18-month rotating schedule. Further, we choose the failure of Lehman Brothers as the starting point (as opposed to August 2007 and the run on the asset-backed commercial paper market) because both CAMELS rating and bank failures start to pick up only notably after the failure of Lehman Brothers, peaking in 2010 (see Figure C.2 in the appendix).

[22] Throughout our regressions, we include the following variables as controls measured as of 2008q2: bank size, equity over assets, time deposits over assets, non-performing loans over total loans, and net income over assets.

assumption is that the assignment of a relatively stricter federal supervisor early in the GFC, right when the broader distress in the U.S. financial sector became very salient, is exogenous to banks that were previously on the rotating schedule. To support the validity of our identification strategy, Table B.4 in the appendix shows that banks scheduled to have FRS-led exams early in the GFC have similar observables as banks scheduled to have state-led exams. In particular, this balance test shows that the two groups of banks are similar along a wide variety of observable characteristics such as pre-GFC CAMELS rating, size, profitability, funding, and capitalization.

The instrument is relevant for exposure to strict supervision, both during the initial phase of the GFC and the later phases. The reason is that initial supervisory scrutiny during the GFC gives rise to a path dependency, in which banks identified as troubled early on are more likely to receive more scrutiny throughout the entire crisis. This notion of path dependency is supported by the evidence in Section 3, which shows that once a bank stops qualifying for the rotating schedule, it is likely to receive additional scrutiny. In that sense, additional scrutiny early in the GFC begets additional scrutiny throughout the GFC. This view, which we reinforce below, is also supported by the evidence that examiners were resource constrained during the GFC (see, e.g., Eisenbach et al., 2022). Thus, banks that were identified as troubled before the spike of bank failures that required supervisors to "triage" were more likely to get additional scrutiny as the GFC played out, as supervisors were naturally prioritizing dealing with banks that had already been labeled as troubled.

4.2 First Stage

We begin by verifying the first stage. We show that banks on the rotating schedule and whose last exam before the GFC was led by a state agency are significantly more likely to be subject to a federal exam early in the GFC. There is no variable reported that indicates whether a given exam is part of a rotating pattern. We thus define a bank as being on the rotating schedule before the GFC if the past four exams had a rotating pattern. For example, a bank whose previous four exams follow the pattern federal/state/federal/state is classified as being on the rotating schedule. Moreover, we restrict the sample to banks that generally qualify for the rotating schedule. Following Agarwal et al. (2014), a bank qualifies for rotation if it is an SMB or NMB that has a CAMELS rating of either 1 or 2, has less than $10 billion in assets, is more than 5 years old, and is not located in either one of the following states: Montana, Rhode Island, Nevada, New Hampshire, South Dakota, South Carolina, Alabama.

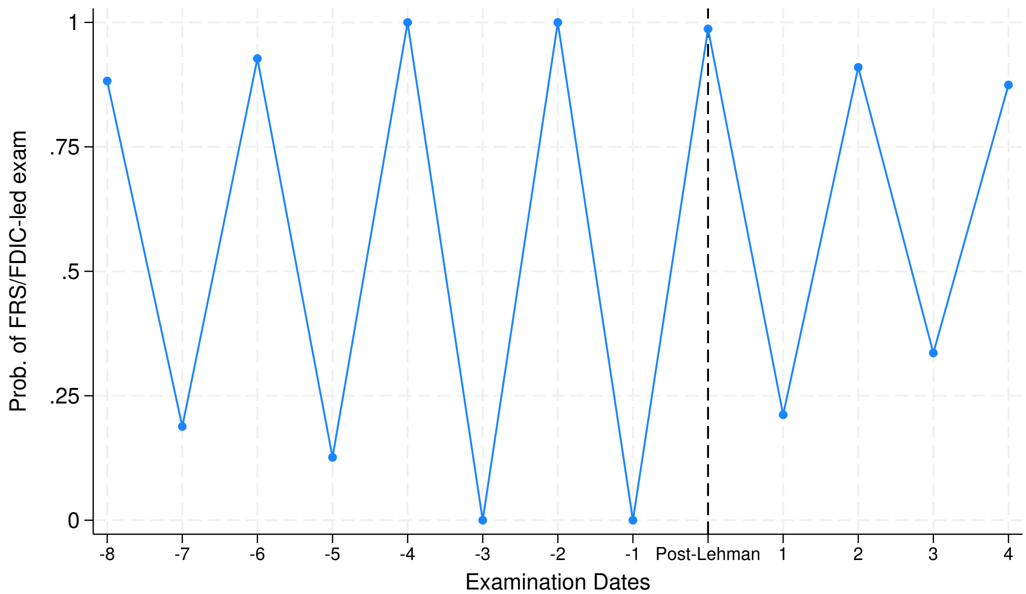
Figure 8 plots the probability of a federal-led (i.e., FDIC-/FRS-led) exam by a bank conditional on the last exam being a state-led exam and the three exams before that exam having a rotating pattern. Thus, the coefficients for the four exams leading up to a federal-led exam during the early GFC are mechanically 0, 1, 0, and 1. As is clear from Figure 8, banks on the rotating schedule before the GFC are highly likely to have a federal-led exam early in the GFC if their last exam before the GFC was state-led. That is, the first stage is very strong. Over time, the rotating pattern becomes less pronounced, in line with banks dropping off the schedule when they disqualify—for instance, after being downgraded. However, the rotating schedule continues to strongly predict which agency will lead the bank examination during the GFC, consistent with path dependence.

4.3 CAMELS Ratings and Bank Examinations

We next trace the evolution of the CAMELS rating throughout the GFC for banks on the rotating schedule before the failure of Lehman Brothers, distinguishing between banks subject to at least one federal exam in the 18 month following the failure of Lehman Brothers (labeled “high supervisory strictness”) and those not subject to a federal-led exam over the same period (labeled “low supervisory strictness”).

Figure 9 plots the within-bank dynamics of the CAMELS rating, normalized to the level in the second quarter of 2008, high scrutiny banks (green line) and low scrutiny

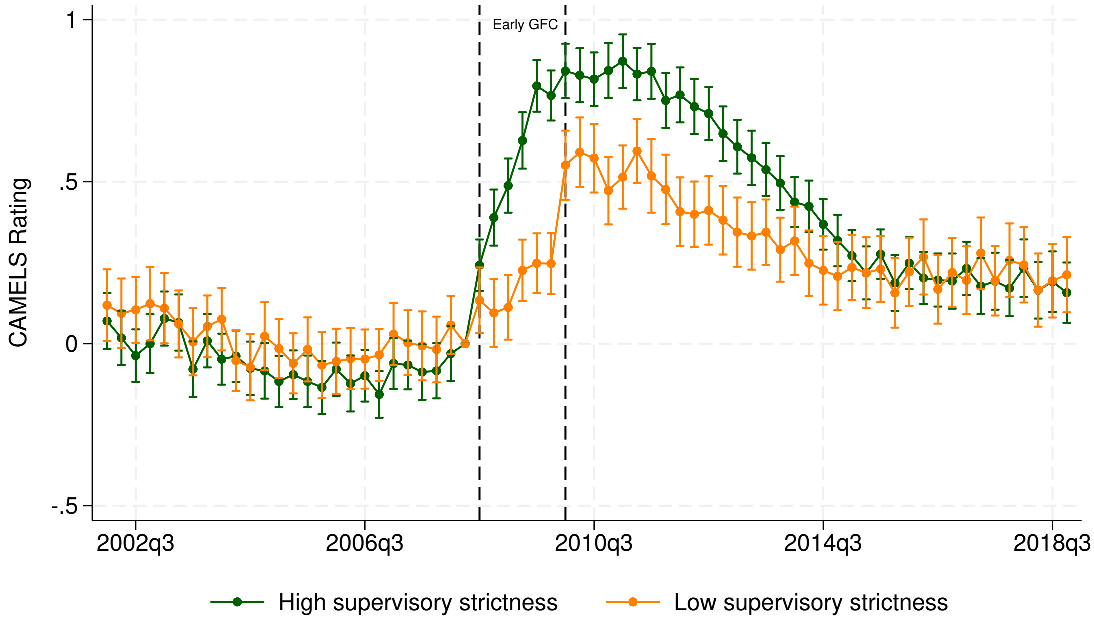
Figure 8: *Probability of Federal Supervisory Exam Early in GFC for Banks on the Rotating Schedule Before the Crisis*



Notes: This figure plots the probability of a federal exam in the 18 months following the failure of Lehman Brother conditional on having the last four consecutive exams before the failure of Lehman rotating between federal and state supervisors. The date "Post-Lehman" is the first exam a bank is subject after the failure of Lehman Brothers.

banks (orange line). There is no differential pre-trend in Figure 9. Right after the second quarter of 2008, however, the average CAMELS rating increases for both groups, in line with rising asset losses and broader distress in the banking sector. Notably, however, banks that were more likely to be subject to a federal-led exam see a faster increase around this time. Thus, in line with federal examiners being relatively tougher, banks that expected a federal exam early in the GFC saw a larger increase in their CAMELS rating and thus the probability of being rated as a troubled bank. This difference across CAMELS ratings narrows in early 2010, exactly when these banks were likely to be scheduled for a federal exam, according to the rotation principle. However, the gap is persistent and disappears only by 2014, after the end of the GFC. Thus, the pattern in Figure 9 indicates that banks subject to additional supervisory scrutiny early in the GFC were also subject to more supervisory scrutiny throughout the GFC, consistent with path dependence based on initial scrutiny.

Figure 9: *CAMELS Rating across Banks Exposed to High and Low Supervisory Strictness during the GFC*



Notes: This figure shows the sequence of coefficients $\{\beta_k\}$ from estimating:

$$y_{bt} = \alpha_b + \sum_{k \neq 2007q4} \beta_k \times \mathbb{I}_{k=t} + \epsilon_{bt},$$

where y_{bt} is bank b's CAMELS rating and α_b is a set of bank fixed effects. We estimate the model for banks subject to both "high" and "low" supervisory strictness separately, defining "high" strictness as having at least one federal agency led exam during the early GFC (in the 18 months following the failure of Lehman Brothers). Standard errors are clustered at the bank level. Error bars represent 95% confidence bands.

The differential effect of supervisory strictness on the CAMELS rating is confirmed when estimating our main regression model Equation (3), instrumenting supervisory strictness with the rotating schedule. Table 5 shows that banks subject to an examination led by a federal agency early in the GFC received a 0.22 higher CAMELS rating, on average, see column (1). The magnitude of this effect is over two times larger than the average effect documented in Agarwal et al. (2014). The differential impact of tougher supervision was thus especially pronounced in the crisis, as one would expect in a setting where the dispersion in underlying bank health widens. Banks exposed to stricter supervision also had a 9 percentage point higher chance of being rated as a troubled bank with a CAMELS of 3 or higher, see column (2).

Table 5: *Effect of Supervisory Strictness on CAMELS and the Number and Length of On-site Examinations during the GFC*

Dependent variable	CAMELS	Prob of CAMELS≥3	No. of FRS/FDIC exam	No. of exams	Length of exams
	(1)	(2)	(3)	(4)	(5)
High supervisory strictness	0.22***	0.095***	0.76***	0.49***	3.90***
	(0.029)	(0.012)	(0.080)	(0.086)	(1.21)
Observations	3604	3604	3604	3604	3575
Mean dep. var	2.05	0.21	2.88	4.59	49.0
R^2	0.22	0.16	0.14	0.11	0.11
State FE	✓	✓	✓	✓	✓
Baseline Controls	✓	✓	✓	✓	✓
First-stage F-stat	6605.2	6605.2	6605.2	6605.2	6556.4

Notes: This table reports results from estimating (3). The outcome, y_b, is either the bank's average CAMELS rating, the probability of being rated with a 3 or higher, the number of total exams, the number of exams led by a federal agency (FDIC or FRS), or the length of the average exam (in days) from 2008q3 through 2013q4. High supervisory strictness$_b$ is a dummy that indicates whether bank b was subject to at least one exam led by a federal agency in the early phase of the GFC, defined as the 18 months following the failure of Lehman Brothers. We instrument whether a bank is subject to a federal exam in the early phase of the GFC with whether its last four previous examinations had rotating exams. Table B.5 shows the same estimation using OLS. The sample is restricted to state-member banks and non-member banks that are supervised by either the Federal Reserve, the FDIC, or state regulatory agencies as well as to banks that qualified for the rotating schedule before 2008q3. Baseline controls are a bank's size, equity over assets, time deposits over assets, non-performing loans over total loans, and net income over assets as of 2008q2. Robust standard errors. *,**, and *** indicate significance at the 10%, 5%, and 1% level, respectively.

Table 5 further shows that banks subject to stricter supervision received more scrutiny from bank examiners. Being identified as a troubled bank leads to additional supervisory scrutiny. In particular, these banks underwent around 0.8 more exams led by a federal agency and 0.5 additional overall exams between fall 2008 and end of 2013. Moreover, banks subject to more scrutiny faced exams that on average lasted around 4 days longer than those subject to low supervisory scrutiny. Importantly, the estimates are robust to applying the instrumental variable approach, addressing the concern that federal examiners select into examining more troubled banks. We note that the OLS and the 2SLS estimates are quite similar in magnitude (see Table B.5 for the OLS estimates). The original selection problem is thus modest in this sample, in line with supervisors having limited resources to take on exams that were not already scheduled beforehand during the early phase of the GFC.

4.4 Financial Statements and Revisions

Being exogenously assigned to a stricter supervisor increases the likelihood of being labeled a troubled bank. It also leads to greater supervisory scrutiny. We now show that this greater scrutiny has important implications for bank financial statements and the likelihood of being subject to enforcement actions, which, in turn, leads to restrictions on bank activities.

Table 6: *Effect of Supervisory Strictness on Revisions of Previously Submitted Call Reports*

Dependent variable	Prob. of downward revision		Prob. of upward revisions	
	Income/Assets	Equity/Assets	LLP/Loans	Charge-Offs/Loans
	(1)	(2)	(3)	(4)
High supervisory strictness	0.013***	0.012***	0.0074***	0.0061***
	(0.0034)	(0.0039)	(0.0026)	(0.0017)
Observations	3607	3607	3607	3607
Mean dep. var	0.047	0.052	0.027	0.014
R^2	0.064	0.069	0.069	0.043
State FE	✓	✓	✓	✓
Baseline Controls	✓	✓	✓	✓
First-stage F-stat	6624.5	6624.5	6624.5	6624.5

Notes: This table reports results from estimating (3), where y_b is a dummy indicating whether a bank b revised the line item indicated in the table header at least once between 2008q3 and 2013q4. In columns (1) and (2) we consider downwards revisions of net income and bank equity. In columns (3) and (4), we consider upwards revisions in loan loss provisions and charge-offs. High supervisory strictness$_b$ is a dummy that indicates whether bank b was subject to at least one exam led by a federal agency in the early phase of the GFC, defined as the 18 months following the failure of Lehman Brothers. The model is estimated using 2SLS. We instrument whether a bank is subject to a federal exam in the early phase of the GFC with whether it was on the rotating schedule in the four previous exams. The sample is restricted to state-member banks and non-member banks that are supervised by either the Federal Reserve, the FDIC, or state regulatory agencies as well as to banks that qualified for the rotating schedule before 2008q3. Baseline controls are a bank's size, equity over assets, time deposits over assets, non-performing loans over total loans, and net income over assets as of 2008q2. Robust standard errors. *,**, and *** indicate significance at the 10%, 5%, and 1% level, respectively.

Table 6 shows results from estimating Equation (3) with an indicator for whether a bank revised selected line items in the Call Report after its initial submission as the outcome variable. Specifically, we ask whether banks that are subject to tougher supervision early in the GFC are more likely to revise their net income or book equity downwards and their loan loss provisions and charge-offs upwards during the GFC. Banks subject to a federal exam early in the GFC have a 1.3 percentage points higher probability of revising their net income downwards (column 1) and a 1.2 percentage points higher probability of revising their equity downwards (column 2). These downward revisions are in part driven by an increase in loan loss provisions, which are 0.7 percentage points more likely to be revised upwards in banks subject to more regulatory scrutiny.

Table 7 examines the effect of stricter supervision on bank income and capitalization. Column (1) in Table 7 shows that banks subject to more supervisory scrutiny have lower average net income (normalized by assets) during the GFC. This finding of a lower net income is in line with supervisors requiring these banks to restate previously submitted financial statements, increasing provisions, and thus increasing the allowances for losses. Indeed, columns (2) and (3) show that banks subject to higher supervisory scrutiny have both higher allowances and provisions.

Taken together, these findings suggest that additional supervisory scrutiny leads banks to recognize losses that would otherwise go unnoticed or be only recognized much later. Huizinga and Laeven (2012) show that banks overinflated asset valuations during the GFC. Stricter supervisors can thus contribute to offsetting inflated asset valuations in a crisis.

4.5 Enforcement, Payouts, and Capitalization

CAMELS downgrades and lower capitalization pave the way for supervisors to take enforcement actions. Consistent with this, Table 8 shows that higher supervisory strictness leads to a 5.7 percentage point higher chance of being subject to some type of enforcement action and a 1.1 percentage point higher chance of being subject to PCA sometime between fall 2008 and end of 2013. Both differences are considerable in terms of their magnitude, given that the overall probability in this sample of enforcement action is 18% and the probability of PCA is only 1.8%.

In line with stricter supervision causing more enforcement actions, we find that banks

Table 7: *Supervisory Scrutiny, Income, and Capitalization.*

Dependent variable	Net income/assets	ALLL/loans	Prov./loans	Cum. Div./Assets$_{2008q2}$	Prob[No div.]	Equity issued	Equity/assets
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
High supervisory strictness	-0.0029***	0.0011***	0.0024***	-0.014***	0.032***	-0.0082	-0.0031***
	(0.00063)	(0.00037)	(0.00051)	(0.0044)	(0.012)	(0.015)	(0.0010)
Observations	3604	3604	3604	3570	3604	3604	3604
Mean dep. var	0.0047	0.018	0.0091	0.11	0.11	0.18	0.091
R^2	0.13	0.26	0.081	0.11	0.084	0.026	0.64
State FE	✓	✓	✓	✓	✓	✓	✓
Baseline Controls	✓	✓	✓	✓	✓	✓	✓
First-stage F-stat	6605.2	6605.2	6605.2	6632.1	6605.2	6605.2	6605.2

Notes: This table reports results from estimating (3), where y_b is outcome variable denoted in the column header. In columns (1) through (3) and (6)- (7), y_b is the average over 2008q3 to 2013q4. In column (4), y_b is the cumulative amount of dividend payouts from 2008q3 to 2013q4 as a share of assets in 2008q2. In column (5), y_b is an indicators that takes the value one if the bank paid no dividend from through 2008q4 to 2013q4. High supervisory strictness$_b$ is a dummy that indicates whether bank b was subject to at least one exam led by a federal agency in the early phase of the GFC, defined as the 18 months following the failure of Lehman Brothers. The model is estimated using 2SLS. We instrument whether a bank is subject to a federal exam in the early phase of the GFC with whether it was having rotating exams in the four previous exams. The sample is restricted to state-member banks and non-member banks that are supervised by either the Federal Reserve, the FDIC, or state regulatory agencies as well as to banks that qualified for the rotating schedule before 2008q3. Baseline controls are a bank's size, equity over assets, time deposits over assets, non-performing loans over total loans, and net income over assets as of 2008q2. Robust standard errors. *,**, and *** indicate significance at the 10%, 5%, and 1% level, respectively.

subject to stricter supervision also pay lower dividends throughout the crisis (Table 7, column 4). The cumulative reduction in dividends amounts to 1.4% of pre-crisis assets, a sizable effect. The probability of paying no dividends from 2008q4 to 2013q4 also rises by 3.2 percentage points. This evidence suggests that stricter supervision can partly counteracted the trend of high dividend payouts during the 2008 crisis (Acharya et al., 2011). At the same time, we find no effect on newly issued equity (Table 7, column 6).

Despite lower dividend payment and thus higher retained earnings, we find that banks subject to more supervisory scrutiny report slightly lower equity-to-assets ratio than those subject to low supervisory scrutiny, see column (7) of Table 7. We interpret this as evidence that the reported book equity in banks that were subjected to more supervisory actions is a better reflection of the market value of the underlying bank assets than in those that were subject to less scrutiny. Therefore, stricter supervision still leads to a higher true level of capitalization through the increase in retained earnings. However, measured capitalization is sufficiently overstated in banks subject to lenient supervision that banks subject to stricter supervision have lower capitalization on paper. This finding

Table 8: *Supervisory Strictness and Enforcement Actions*

Dependent variable	Enforcement	PCA
	(1)	(2)
High supervisory strictness	0.057***	0.011**
	(0.015)	(0.0050)
Observations	3607	3607
Mean dep. var	0.18	0.018
R^2	0.071	0.021
State FE	✓	✓
Baseline Controls	✓	✓
First-stage F-stat	6624.5	6624.5

Notes: This table reports results from estimating (3), where y_b is indicates whether a bank was subject to any public enforcement action or whether the bank was subject to Prompt Corrective Action (PCA) at any point in time between 2008q3 and 2013q4. High supervisory strictness$_b$ is a dummy that indicates whether bank b was subject to at least one exam led by a federal agency in the early phase of the GFC, defined as the 18 months following the failure of Lehman Brothers. The model is estimated using 2SLS. We instrument whether a bank is subject to a federal exam in the early phase of the GFC with whether it was having rotating exams in the four previous exams. The sample is restricted to state-member banks and non-member banks that are supervised by either the Federal Reserve, the FDIC, or state regulatory agencies as well as to banks that qualified for the rotating schedule before 2008q3. Baseline controls are a bank's size, equity over assets, time deposits over assets, non-performing loans over total loans, and net income over assets as of 2008q2. Robust standard errors. *,**, and *** indicate significance at the 10%, 5%, and 1% level, respectively.

is in line with the pattern documented in James (1991), who finds that failing banks with more book equity at failure tend to have higher losses due to previously unreported losses.

4.6 Bank Closures and the Cost of Failure

Table 9 studies the effect of stricter supervision in the GFC on the probability of bank failure, the speed of failure, and cost of failures to the FDIC's DIF. We find that banks subject to tougher supervision were 3.2 percentage points more likely to fail from 2008 through 2013. This is a large difference, considering that the unconditional failure rate is 3.4%. Thus, more supervisory scrutiny leads to *more* bank failures. This finding is strikingly in contrast to an often-held view that supervisory actions should reduce the

Table 9: *Failure, Speed of Failure, and Costs for the FDIC Insurance Fund.*

Dependent variable	Probability of failure		Speed of failure (days)		DIF costs/assets	
	OLS	2SLS	OLS	2SLS	OLS	2SLS
	(1)	(2)	(3)	(4)	(5)	(6)
High supervisory strictness	0.025***	0.032***	-279.3***	-432.4**	-0.049*	-0.092*
	(0.0045)	(0.0066)	(103.0)	(186.2)	(0.029)	(0.049)
Observations	3610	3610	116	116	116	116
Mean dep. var	0.034	0.034	962.1	962.1	0.22	0.22
R^2	0.12	0.053	0.34	0.15	0.52	0.15
State FE	✓	✓	✓	✓	✓	✓
Baseline Controls	✓	✓	✓	✓	✓	✓
First-stage F-stat		6615.1		16.3		16.3

Notes: This table reports results from estimating (3), where y_b is a dummy whether the bank failed between 2008q3 and 2013q4, the speed of failure (how many days passed between the failure of Lehman Brothers and failure of bank b), and the cost the FDIC incurred as a share of the failed banks assets at failure. High supervisory strictness$_b$ is a dummy that indicates whether bank b was subject to at least one exam led by a federal agency in the early phase of the GFC, defined as the 18 months following the failure of Lehman Brothers. We provide both OLS and 2SLS estimates. In the 2SLS model, we instrument whether a bank is subject to a federal exam in the early phase of the GFC with whether it was having rotating exams in the four previous exams. The sample is restricted to state-member banks and non-member banks that are supervised by either the Federal Reserve, the FDIC, or state regulatory agencies as well as to banks that qualified for the rotating schedule before 2008q3. Baseline controls are a bank's size, equity over assets, time deposits over assets, non-performing loans over total loans, and net income over assets as of 2008q2. Robust standard errors. *,**, and *** indicate significance at the 10%, 5%, and 1% level, respectively.

chance of bank failures. While stricter supervision could reduce the *ex ante* risk of failures, this evidence indicates that an important role of supervision in the wake of a large asset shock when many banks are likely to be insolvent is to impose discipline and ultimately close insolvent banks.

In addition to increasing the likelihood of failure, stricter supervision increases the speed of failure. Columns (3) and (4) of Table 9 show that banks subject to tougher supervision close their operations more than one year earlier than those that are subject to lenient supervision. Stricter supervision thus resolves banks in a timelier fashion.

Stricter supervision has the benefit of leading to a lower cost of failure for the FDIC, conditional on failure. As noted above, the FDIC tends to realize losses when banks fail. These losses can be either due to losses from the resolved bank's investments or inefficiency associated with the process of selling failed banks (Granja, 2013; Granja et al., 2017; Allen et al., 2023a). Columns (5) and (6) of Table 9 show that, conditional on failing, banks that are subject to tougher supervision early in the GFC have a cost-to-assets ratio that is around 9 percentage points lower than those subject to lenient supervision. This effect is considerable given the average cost to the FDIC is 22% of a failing bank's assets. Figure C.4 in the Appendix shows the result visually: the distribution of the cost of failure is shifted to the left for banks subject to strict supervision.

Why are failures for banks exposed to strict supervision less costly for the FDIC? There are at least four explanations. First, strict supervision leads banks to pay out fewer dividends and increase retained earnings. If the distribution of true bank capitalization is shifted rightward by stricter supervision, then failures will also involve less undercapitalized banks, for a given failure rate, leading to a lower cost of failure for the FDIC. This mechanism can explain a non-trivial share of the effect, given the finding that stricter supervision led to a 1.4% increase in retained earnings-to-assets for the average bank.

Second, and related to the first point, stricter supervisors may close healthier banks on the margin. The fact that strict supervisors increase bank capitalization and close more banks suggests the marginal bank closed by a strict supervisor will be better capitalized than one closed by a lenient supervisor.

Third, closing banks faster may reduce the scope for value destroying actions such as gambling for resurrection. Table B.7 includes the speed of failure as a control variable in the cost of failure regression. Under the strong assumption that the speed of failure is exogenous in this regression, we can perform a simple mediation analysis (MacKinnon, 2012).[23] Given that strict supervision reduces the time to failure by 432.4 days, and using that a reduction in time to failure by one day reduces the cost of failure by 0.000073 as a share of assets (Table B.7), we have that the effect through a reduction in the time to failure can account for 3.2 pp ($= 0.000073 \times 432.4$) of the overall reduction in the cost of failure of 9.2% of assets. Thus, based on admittedly strong assumptions, the acceleration

[23]This is a strong assumption, as banks exposed to larger asset losses are likely to have a higher cost of failure and may also fail faster.

of failure can account for about one-third of the reduction in the cost of failure induced by strict supervision. Given that the banks with larger losses are likely to fail faster, we can view this as an upper bound on the role of this channel. These estimates provide causal evidence for arguments that failure to resolve insolvent banks quickly increased the cost of bank failures to the FDIC (e.g., Cole and White, 2017).

One factor potentially offsetting the benefit of closing banks faster is that these banks may be sold into less liquid markets, as the market for failing banks may be less liquid early in a crisis. For example, Figure C.5 uses data from Allen et al. (2023a) to show that the average number of bidders and bids for failed banks was two-to-three times lower in 2009-2011 than 2012-2014. Therefore, while resolving banks more quickly likely reduced the cost of failure, this may have been offset by reduced liquidity for these banks in the FDIC's failed bank auctions.

Finally, stricter supervision may lower the cost of failure through more accurate financial statements. This, in turn, can increase the efficiency of the FDIC's Purchase and Assumption process.

Overall, the reduction in the cost of failure from stricter supervision is thus likely to be driven by a combination of these banks having higher true capitalization, being healthier on the margin, undertaking less value destroying actions by being closed more quickly, and having a more efficient resolution process due to factors such as improved accounting and increased transparency.

4.7 Heterogeneity across Exposure to the Housing Market Crash

The impact of stricter supervision should be more pronounced for banks with greater exposure to asset losses. A larger asset shock implies a greater role for supervisors to close insolvent banks in a time manner. To test this idea, we explore whether the findings presented in Table 9 are driven by banks that were more exposed to the decline in the local housing market. To measure a bank's exposure to the housing crash, we combine information on county-level house price growth from 2005 through 2008 and the geographic footprint of banks proxied by county-level deposit market shares.[24] Specifically, we define the the exposure of bank b as the deposit market share weighted average of housing price growth in each county c:

$$\text{GFC Exposure}_b = \sum_c \text{Deposit market share}_{bc} \times \Delta\text{Housing Price Index}_{c,2005-2008}. \tag{4}$$

We then split our sample into banks with above- and below-median exposure to the GFC and re-estimate Equation (3) for both samples.

Table 10: *Failure, Speed of Failure, and FDIC Insurance Fund Costs by GFC Exposure.*

Dependent variable	Probability of failure		Speed of failure (days)		DIF costs/assets	
GFC Exposure	High	Low	High	Low	High	Low
	(1)	(2)	(3)	(4)	(5)	(6)
High supervisory strictness	0.043***	0.021***	-442.9**	1145.3	-0.087*	-0.56
	(0.015)	(0.0058)	(173.6)	(3271.4)	(0.047)	(1.15)
Observations	1516	2090	80	31	80	31
Mean dep. var	0.057	0.017	970.1	954.7	0.22	0.23
R^2	0.071	0.031	0.17	0.22	0.086	0.46
State FE	✓	✓	✓	✓	✓	✓
Baseline Controls	✓	✓	✓	✓	✓	✓
First-stage F-stat	1532.7	6044.7	19.1	1.07	19.1	1.07

Notes: This table reports results from estimating (3), where y_b is either a dummy whether the bank failed between 2008q3 and 2013q4, the speed of failure (how many days passed between the failure of Lehman Brothers and failure of bank b), or the cost the FDIC incurred as a share of the failed banks assets at failure. High supervisory strictness$_b$ is a dummy that indicates whether bank b was subject to at least one exam led by a federal agency in the early phase of the GFC, defined as the 18 months following the failure of Lehman Brothers. We split the sample into banks with an above or below median exposure to the GFC, with the exposure being calculated by averaging across the housing price growth in each county c and using deposit market shares as weights, as defined in (4). High supervisory strictness$_b$ is instrumented with whether it a bank was on the rotating schedule in the four previous exams. The sample is restricted to state-member banks and non-member banks that are supervised by either the Federal Reserve, the FDIC, or state regulatory agencies as well as to banks that qualified for the rotating schedule before 2008q3. Baseline controls are a bank's size, equity over assets, time deposits over assets, non-performing loans over total loans, and net income over assets as of 2008q2. Robust standard errors. *,**, and *** indicate significance at the 10%, 5%, and 1% level, respectively.

[24] House price growth from 2005 to 2008 captures the drop in house prices in the early phase of the Great Recession. Mian and Sufi (2009) and Mian and Sufi (2014) documents that this shock had strong effects on local defaults, consumption, and employment.

Table 10 shows that the impact of supervisory strictness on the likelihood, speed, and cost of failure is primarily driven by banks that have a high exposure to the GFC. That is, the result is strongest where one would it expect it to be. For instance, the effect of high supervisory strictness on the probability of failure is 4.3 percentage points in banks with a high exposure, but only 2.1 percentage points in low exposure banks. Moreover, the effect of high supervisory strictness on the speed of failure and the cost of the FDIC insurance fund is largely driven by banks with a high exposure to the housing crash, as evidenced by the almost identical coefficients across both the estimation using the entire sample in Table 9 and using only the highly exposed banks in Table 10. Given that the estimation in columns (3) through (6) is conditional on failure and failure less likely for low exposure banks, the instrument becomes too weak due to the small sample size in columns (4) and (6).

4.8 Lending and Inputs to Banking

Stricter supervision, by identifying losses and taking enforcement actions in troubled banks, may restrict banking activities. This can lead to less lending and reduced inputs to the production of banking services, such as a reduction in branches and employees. Table 10 examines the effect of stricter supervision on bank lending and the inputs to banking. In the short run, banks exposed to stricter supervision see a significantly larger contraction in total loans, real estate loans, and C&I loans from fall 2008 through end of 2013. For example, stricter supervision leads to a 5.8% reduction in total loans.

There are at least two explanations for the decline in lending. First, greater loss recognition reduces bank equity, tightening capital constraints. This reduces the ability to lend, exacerbating the credit crunch. A second explanation is that stricter supervision reduces evergreening of loans that are unlikely to be repaid. In this sense, reduced lending reflects the recognition of losses, accelerating a necessary cleansing during a recession (Caballero et al., 2008).

It is challenging to fully disentangle the role of these two explanations using balance sheet data from the Call Reports. We believe both are likely to be operative. The reduced evergreening channel is likely relevant, given that stricter supervisors force banks to recognize more losses. However, we also believe that a reduction in credit availability also plays a role. For one, we observe reduced lending across all loan types, including C&I loans, where defaults were lower compared to real estate loans.

Another piece of evidence that supports reduced lending is that stricter supervision led to a contraction in the inputs to banking. In particular, Table 10 shows that banks exposed to stricter supervision reduced branches by 3% and employees by 2% from 2008 to 2013. Several studies document distance constraints in lending (e.g. Degryse and Ongena, 2005; Mian, 2006). Moreover, branch closures often reduce credit access for local borrowers, with real effects (Nguyen, 2019; Amberg and Becker, 2024; Ranish et al., 2024). Barring an increase in bank efficiency that would allow the bank to maintain information production and lending with reduced inputs, the evidence points to a reduction in credit availability.

The evidence thus points to a policy tradeoff. On the one hand, stricter supervision leads to increased loss recognition, faster bank closures, and a lower cost of closures. Stricter supervision can thus mitigate the adverse consequences of forbearance. Regulatory forbearance is often described as a key culprit behind the high cost of bank failures during the S&L Crisis (Kane, 1989b) and lower productivity in the aftermath of the Japanese and European financial crises (Caballero et al., 2008; Blattner et al., 2023). Indeed, the FDIC Improvement Act of 1991 sought to reduce forbearance and expedite the closure of insolvent banks to reduce the cost of failures. On the other hand, strict supervision poses the risk of making policy overly procyclical, reducing credit availability and exacerbating a credit crunch. The closure of bank branches or of an insolvent bank, while potentially optimal, may nevertheless entail a loss in the local information-producing services of banking that may be costly to replace. In a crisis, when credit and demand may be inefficiently low, the reduction in lending may exacerbate the downturn in the short run.

Table 11: *Supervisory Strictness, Lending, and Bank Branches and Employment*

Panel A: Growth from 2008q2 to 2013q4

Dependent variable	ΔLoans	Δ RE loans	Δ C&I loans	Δ CRE loans	Δ Branches	Δ Employees
Period	2008q2-2013q4					
	(1)	(2)	(3)	(4)	(5)	(6)
High supervisory strictness	-0.058***	-0.046**	-0.065**	-0.068**	-0.031***	-0.021**
	(0.019)	(0.022)	(0.031)	(0.030)	(0.010)	(0.010)
Observations	3099	3087	3074	3061	3034	2960
Mean dep. var	0.14	0.18	0.027	0.11	0.060	0.050
R^2	0.041	0.032	0.023	0.030	0.012	0.022
State FE	✓	✓	✓	✓	✓	✓
Baseline Controls	✓	✓	✓	✓	✓	✓
First-stage F-stat	6009.6	5949.0	6077.7	5862.1	5981.5	6104.9

Panel B: Growth from 2008q2 to 2018q4

Dependent variable	ΔLoans	Δ RE loans	Δ C&I loans	Δ CRE loans	Δ Branches	Δ Employees
Period	2008q2-2018q4					
	(1)	(2)	(3)	(4)	(5)	(6)
High supervisory strictness	-0.0046	0.0078	0.0017	-0.035	-0.034*	-0.0058
	(0.030)	(0.032)	(0.041)	(0.044)	(0.020)	(0.022)
Observations	2562	2550	2539	2526	2562	2560
Mean dep. var	0.49	0.55	0.37	0.51	0.17	0.17
R^2	0.039	0.017	0.020	0.016	0.010	0.0090
State FE	✓	✓	✓	✓	✓	✓
Baseline Controls	✓	✓	✓	✓	✓	✓
First-stage F-stat	5838.9	5806.7	5833.6	5702.7	5838.9	5860.7

Notes: This table reports results from estimating (3), where y_b is the outcome in the column header. High supervisory strictness$_b$ is a dummy that indicates whether bank b was subject to at least one exam led by a federal agency in the early phase of the GFC, defined as the 18 months following the failure of Lehman Brothers. The model is estimated using 2SLS. We instrument whether a bank is subject to a federal exam in the early phase of the GFC with whether it was having rotating exams in the four previous exams. The sample is restricted to state-member banks and non-member banks that are supervised by either the Federal Reserve, the FDIC, or state regulatory agencies as well as to banks that qualified for the rotating schedule before 2008q3. Baseline controls are a bank's size, equity over assets, time deposits over assets, non-performing loans over total loans, and net income over assets as of 2008q2. Robust standard errors. *,**, and *** indicate significance at the 10%, 5%, and 1% level, respectively.

5 Conclusion

We study the role of banking supervision in monitoring and imposing discipline on troubled banks. Our analysis leverages confidential supervisory data for the U.S. banking system from 2000 to 2023. We find that supervisory ratings have strong predictability for future bank failures in the near term. Despite notable exceptions, supervisors are rarely blindsided by bank failures. Supervisors play a central role in closing insolvent banks, especially in the absence of depositor market discipline due to deposit insurance. Outside of failure, an important role of supervision is to ensure that banks' financial statements are accurate. Supervisors also take enforcement actions in troubled banks, though prompt corrective action, the most stringent action, is uncommon.

To establish causality, we estimate the impact of stricter supervision using exogenous variation in the assignment of supervisors during the onset of the 2008 Global Financial Crisis. We find that stricter supervisors are more likely to downgrade bank CAMELS ratings and increase scrutiny on troubled banks. This translates into faster recognition of loan losses, increased enforcement actions, and reduced dividend payouts. Stricter supervisors also closed more banks and did so sooner in the GFC, leading to a lower cost of failures to the DIF. These benefits came at the expense of a reduction in the inputs to banking and reduced credit, highlighting a tradeoff of strict supervision. Thus, our paper highlights the importance of banking supervision in overseeing troubled banks, which is complementary to the effect of banking supervision on lending (see, e.g., Granja et al., 2017) or risk-taking (see, e.g., Hirtle et al., 2020; Kandrac and Schlusche, 2020) highlighted in previous work.

There are important caveats to keep in mind when interpreting our findings. First, most bank failures we consider are due to realized credit risk rather than interest rate risk (Gopalan and Granja, 2024), which has been receiving more attention recently. Second, most banks in our sample are small community banks, and the role of supervision may be different in large and complex banking institutions (Hirtle et al., 2020; Eisenbach et al., 2022). Third, while our findings have the implication that raising supervisory scrutiny likely ensures more accurate financial reporting as well as enforcement actions and bank closures when legally required, it is difficult to draw normative implications. That is, we cannot assess whether the overall level of supervisory scrutiny in the U.S. banking system is too aggressive or too lenient and whether too many or too few banks are being closed as a consequence of supervisory actions. Our findings also point to potential short-term economic costs of excessive strictness in a crisis. Fully evaluating the welfare costs and benefits of the stance of supervision is an important question for future research.

Keeping these caveats in mind, we believe our paper makes an important contribution by documenting the potentially important role of supervision in ensuring that failing banks are identified as such and closed. If conducted efficiently, this function of banking supervision is likely valuable and in the broader public interest. In particular, in a banking system in which banks are at least partially insulated from market discipline due to deposit insurance, supervisory actions may be required in order to impose discipline on failing banks. This discipline, as described by Diamond and Rajan (2001), ensures that failing banks promptly recognize losses on past investments, thus preventing them from "kicking the can down the road." This finding is especially relevant in light of ongoing debates on how to reform and consolidate the supervisory and regulatory landscape.

Based on the evidence we provide, our preferred analogy for how supervision operates in failing banks is that of a city health official who conducts food inspections in restaurants and provides grades and ratings on the hygiene standards. Such health officials are unable to advise a restaurant with a failing business model on how to turn the business around (as much as a supervisor may face difficulty in assessing what types of lending decisions are more or less profitable). Rather, they can simply provide guidance to customers on whether it is safe to eat at an establishment or not. Moreover, if it becomes unsafe to dine, they can mandate a temporary or permanent closing. Such closings can help safeguard public safety and thereby contribute to the common good.

References

Acharya, V. V., L. Borchert, M. Jager, and S. Steffen (2021, 01). Kicking the can down the road: Government interventions in the european banking sector. *The Review of Financial Studies 34*(9), 4090–4131.

Acharya, V. V., I. Gujral, N. Kulkarni, and H. S. Shin (2011). Dividends and bank capital in the financial crisis of 2007-2009. Technical report, National Bureau of Economic Research.

Agarwal, S., D. Lucca, A. Seru, and F. Trebbi (2014). Inconsistent Regulators: Evidence from Banking *. *The Quarterly Journal of Economics 129*(2), 889–938.

Agarwal, S., B. C. Morais, A. Seru, and K. Shue (2024). Noisy experts? discretion in regulation. Working Paper 32344, National Bureau of Economic Research.

Allen, J., R. Clark, B. Hickman, and E. Richert (2023a, 06). Resolving failed banks: Uncertainty, multiple bidding and auction design. *The Review of Economic Studies 91*(3), 1201–1242.

Allen, J. J., R. Clark, B. Hickman, and E. Richert (2023b). Banking fragility and resolution costs. SSRN Working Paper, posted June 20, 2023. Available at SSRN: https://ssrn.com/abstract=4434353.

Altavilla, C., M. Boucinha, J.-L. Peydró, and F. Smets (2020). Banking supervision, monetary policy and risk-taking: Big data evidence from 15 credit registers. Technical report, Kiel, Hamburg.

Amberg, N. and B. Becker (2024). Banking without branches. Technical report, Sveriges Riksbank Working Paper Series.

Badertscher, B. A., J. J. Burks, and P. D. Easton (2018). The market reaction to bank regulatory reports. *Review of Accounting Studies 23*(2), 686–731.

Baer, G. (2024, July). The bank examination problem, and how to fix it. Technical report, Bank Policy Institute.

Baron, M., E. Verner, and W. Xiong (2021). Banking crises without panics. *The Quarterly Journal of Economics 136*(1), 51–113.

Baron, M. and W. Xiong (2017). Credit expansion and neglected crash risk. *The Quarterly Journal of Economics 132*(2), 713–764.

Barr, M. (2023). Review of the federal reserve's supervision and regulation of silicon valley bank.

Barth, J. R., G. Caprio, and R. Levine (2012). *Guardians of Finance: Making Regulators Work for Us*. The MIT Press.

Ben-David, I., A. A. Palvia, and R. M. Stulz (2019). Do distressed banks really gamble for resurrection? Working Paper 25794, National Bureau of Economic Research.

Berger, A. N. and S. M. Davies (1998). The information content of bank examinations. *Journal of Financial Services Research 14*(2), 117–144.

Berger, A. N., S. M. Davies, and M. 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.

Bernanke, B. S., T. F. Geithner, and H. M. Paulson Jr (2019). *Firefighting: The Financial Crisis and Its Lessons*. Penguin.

Bernanke, B. S., C. S. Lown, and B. M. Friedman (1991). The credit crunch. *Brookings papers on economic activity 1991*(2), 205–247.

Bessent, S. (2025, April). Remarks by secretary of the treasury scott bessent before the american bankers association. https://home.treasury.gov/news/press-releases/sb0078. U.S. Department of the Treasury, April 9, 2025.

Blattner, L., L. Farinha, and F. Rebelo (2023). When losses turn into loans: The cost of weak banks. *American Economic Review 113*(6), 1600–1641.

Board of Governors (2005). *The Federal Reserve System: Purposes and Functions*. Board of Governors of the Federal Reserve System.

Board of Governors (2023). Review of the federal reserve's supervision and regulation of silicon valley bank. Technical Report April 28, 2023, Federal Reserve.

Bonfim, D., G. Cerqueiro, H. Degryse, and S. Ongena (2023). On-site inspecting zombie lending. *Management Science 69*(5), 2547–2567.

Brown, C. O. and I. S. Dinç (2005). The politics of bank failures: Evidence from emerging markets. *The Quarterly Journal of Economics 120*(4), 1413–1444.

Caballero, R. J., T. Hoshi, and A. K. Kashyap (2008, December). Zombie lending and depressed restructuring in japan. *American Economic Review 98*(5), 1943–77.

Calomiris, C. W. and M. S. Jaremski (2019, April). Stealing deposits: Deposit insurance, risk-taking, and the removal of market discipline in early 20th-century banks. *Journal of Finance 74*(2), 711–754.

Cole, R. and J. Gunther (1998, April). Predicting Bank Failures: A Comparison of On- and Off-Site Monitoring Systems. *Journal of Financial Services Research 13*(2), 103–117.

Cole, R. and L. White (2012, October). Déjà Vu All Over Again: The Causes of U.S. Commercial Bank Failures This Time Around. *Journal of Financial Services Research 42*(1), 5–29.

Cole, R. A. and L. J. White (2017). When time is not on our side: The costs of regulatory forbearance in the closure of insolvent banks. *Journal of Banking & Finance 80*, 235–249.

Correia, S. A., S. Luck, and E. Verner (2024, September). "failing banks". Working Paper 32907, National Bureau of Economic Research.

Costello, A. M., J. Granja, and J. Weber (2019, June). Do Strict Regulators Increase the Transparency of Banks? *Journal of Accounting Research 57*(3), 603–637.

Cucic, D., R. Iyer, S. Kokas, J. Peydró, and S. Pica (2024). Distortive effects of deposit insurance: Administrative evidence from deposit and loan accounts. Technical report. Draft presented at various research conferences; please cite authors' permissions.

Curry, T. J., J. Coburn, and L. Montgomery (1999). Financially distressed banks: How effective are enforcement actions in the supervision process? *FDIC Banking Review 12*, 1–18.

Dahl, D., J. O'Keefe, and G. Hanweck (1998). The influence of examiners and auditors on loan-loss recognition. *FDIC Banking Review 11*, 10–25.

Degryse, H. and S. Ongena (2005). Distance, lending relationships, and competition. *The Journal of Finance 60*(1), 231–266.

Diamond, D. W. and R. G. Rajan (2001). Liquidity risk, liquidity creation, and financial fragility: A theory of banking. *Journal of Political Economy 109*(2), 287–327.

Drehmann, M. and M. Juselius (2014). Evaluating early warning indicators of banking crises: Satisfying policy requirements. *International Journal of Forecasting 30*(3), 759–780.

Eisenbach, T., A. Haughwout, B. Hirtle, A. Kovner, D. Lucca, and M. Plosser (2017). Supervising large, complex financial institutions: what do supervisors do? *Economic Policy Review 23*(1), 57–77.

Eisenbach, T. M., D. O. Lucca, and R. M. Townsend (2022). Resource allocation in bank supervision: Trade-offs and outcomes. *The Journal of Finance 77*(3), 1685–1736.

Fahlenbrach, R., R. Prilmeier, and R. M. Stulz (2018). Why does fast loan growth predict poor performance for banks? *The Review of Financial Studies 31*(3), 1014–1063.

FDIC (2023). Options for deposit insurance reform. https://www.fdic.gov/analysis/options-deposit-insurance-reforms/index.html.

Federal Reserve Bank of Dallas (1979, November). Uniform rating system, circular no. 79-191. Technical Report Circular No. 79-191, Federal Reserve Bank of Dallas, Dallas, Texas. "Uniform Rating System," issued to all state member banks in the Eleventh Federal Reserve District.

Financial Crisis Inquiry Commission (2011). *The Financial Crisis Inquiry Report*. Washington D.C.: Printing Office of the US Government.

Gopalan, Y. and J. Granja (2024). How (in)effective was bank supervision during the 2022 monetary tightening? Working Paper.

Granja, J. (2013). The relation between bank resolutions and information environment: Evidence from the auctions for failed banks. *Journal of Accounting Research 51*(5), 1031–1070.

Granja, J. (2018). Disclosure regulation in the commercial banking industry: Lessons from the national banking era. *Journal of Accounting Research 56*(1), 173–216.

Granja, J. and C. Leuz (2024). The death of a regulator: Strict supervision, bank lending, and business activity. *Journal of Financial Economics 158*, 103871.

Granja, J., G. Matvos, and A. Seru (2017). Selling failed banks. *The Journal of Finance 72*(4), 1723–1784.

Greenwood, R., S. G. Hanson, A. Shleifer, and J. A. Sørensen (2022). Predictable financial crises. *The Journal of Finance 77*(2), 863–921.

Gunther, J. W. and R. R. Moore (2002, None). Loss underreporting and the auditing role of bank exams. *Conference Series ; [Proceedings] None*(None), None.

Hirtle, B. and A. Kovner (2022). Bank supervision. *Annual Review of Financial Economics 14*(Volume 14, 2022), 39–56.

Hirtle, B., A. Kovner, and M. Plosser (2020). The impact of supervision on bank performance. *The Journal of Finance 75*(5), 2765–2808.

Hirtle, B. and J. A. Lopez (1999). Supervisory information and the frequency of bank examinations. *Economic Policy Review 5*(Apr), 1–20.

Huizinga, H. and L. Laeven (2012). Bank valuation and accounting discretion during a financial crisis. *Journal of financial economics 106*(3), 614–634.

Ivanov, I. and S. Karolyi (2023). Fighting failure: The persistent real effects of resolving distressed banks. Working Paper.

Ivanov, I. T., B. Ranish, and J. Wang (2023, March). Banks' Strategic Responses to Supervisory Coverage: Evidence from a Natural Experiment. *Journal of Money, Credit and Banking 55*(2-3), 503–530.

Ivanov, I. T. and J. Z. Wang (2024, May). Bank Supervision and Corporate Credit Supply. *Management Science 70*(5), 3338–3361.

Iyer, R., T. Lærkholm Jensen, N. Johannesen, and A. Sheridan (2019, 03). The distortive effects of too big to fail: Evidence from the danish market for retail deposits. *The Review of Financial Studies 32*(12), 4653–4695.

James, C. (1991). The losses realized in bank failures. *The Journal of Finance 46*(4), 1223–1242.

Kandrac, J. and B. Schlusche (2020, 09). The effect of bank supervision and examination on risk taking: Evidence from a natural experiment. *The Review of Financial Studies 34*(6), 3181–3212.

Kane, E. J. (1989a, December). The high cost of incompletely funding the fslic shortage of explicit capital. *Journal of Economic Perspectives 3*(4), 31–47.

Kane, E. J. (1989b). The s&l insurance mess: How did it happen?

Kareken, J. H. and N. Wallace (1978). Deposit insurance and bank regulation: A partial-equilibrium exposition. *Journal of business*, 413–438.

Kroszner, R. S. and P. E. Strahan (1996). Regulatory incentives and the thrift crisis: Dividends, mutual-to-stock conversions, and financial distress. *The Journal of Finance 51*(4), 1285–1319.

Lucca, D., A. Seru, and F. Trebbi (2014). The revolving door and worker flows in banking regulation. *Journal of Monetary Economics 65*(C), 17–32.

MacKinnon, D. (2012). *Introduction to statistical mediation analysis*. Routledge.

Martin, C., M. Puri, and A. Ufier (2023). Deposit inflows and outflows in failing banks: The role of deposit insurance. *Journal of Finance (Forthcoming)*.

Mian, A. (2006). Distance constraints: The limits of foreign lending in poor economies. *The Journal of Finance 61*(3), 1465–1505.

Mian, A. and A. Sufi (2009). The consequences of mortgage credit expansion: Evidence from the us mortgage default crisis. *The Quarterly journal of economics 124*(4), 1449–1496.

Mian, A. and A. Sufi (2014). What explains the 2007–2009 drop in employment? *Econometrica 82*(6), 2197–2223.

Müller, K. and E. Verner (2024). Credit allocation and macroeconomic fluctuations. *Review of Economic Studies 91*(6), 3645–3676.

Nguyen, H.-L. Q. (2019). Are credit markets still local? evidence from bank branch closings. *American Economic Journal: Applied Economics 11*(1), 1–32.

Oshinsky, R. and V. Olin (2005). Troubled banks: Why don't they all fail? *FDIC Banking Review 18*.

Passalacqua, A., P. Angelini, F. Lotti, and G. Soggia (2021). The Real Effects of Bank Supervision: Evidence from On-Site Bank Inspections. Temi di discussione (Economic working papers) 1349, Bank of Italy, Economic Research and International Relations Area.

Ranish, B., A. Stella, and J. Zhang (2024). Out of sight, out of mind: Nearby branch closures and small business growth. Finance and Economics Discussion Series (FEDS) Paper 2024-071, Board of Governors of the Federal Reserve System.

Schularick, M. and A. M. Taylor (2012). Credit booms gone bust: monetary policy, leverage cycles, and financial crises, 1870–2008. *American Economic Review 102*(2), 1029–1061.

Seru, A. (2025, January). Wall street regulation needs a rethink under donald trump. *Financial Times*. Opinion / financial regulation.

Walter, J. R. (2004). Closing troubled banks: How the process works. *FRB Richmond Economic Quarterly 90*(1), 51–68.

Wheelock, D. C. and P. W. Wilson (2000). Why do banks disappear? the determinants of us bank failures and acquisitions. *Review of Economics and Statistics 82*(1), 127–138.

Zingales, L. (2023, March). Investigate the bank failures. *City Journal*. Eye on the News / Economy, Finance, and Budgets.

Supervising Failing Banks

Online Appendix

Sergio Correia, Stephan Luck, and Emil Verner*

*Correia: Federal Reserve Bank of Richmond, sergio.correia@rich.frb.org; Luck: Federal Reserve Bank of New York, stephan.luck@ny.frb.org; Verner: MIT Sloan and NBER, everner@mit.edu.

A Data Appendix

A.1 CAMELS Rating System

In the following we describe the different CAMELS rating categories in detail. For more background, see the FDIC examination manual here.

- Composite 1: Financial institutions in this group are sound in every respect and generally have components rated 1 or 2. Any weaknesses are minor and can be handled in a routine manner by the board of directors and management. These financial institutions are the most capable of withstanding the vagaries of business conditions and are resistant to outside influences such as economic instability in their trade area. These financial institutions are in substantial compliance with laws and regulations. As a result, these financial institutions exhibit the strongest performance and risk management practices relative to the institution's size, complexity, and risk profile, and give no cause for supervisory concern

- Composite 2: Financial institutions in this group are fundamentally sound. For a financial institution to receive this rating, generally no component rating should be more severe than 3. Only moderate weaknesses are present and are well within the board of directors' and management's capabilities and willingness to correct. These financial institutions are stable and are capable of withstanding business fluctuations. These financial institutions are in substantial compliance with laws and regulations. Overall risk management practices are satisfactory relative to the institution's size, complexity, and risk profile. There are no material supervisory concerns and, as a result, the supervisory response is informal and limited.

- Composite 3: Financial institutions in this group exhibit some degree of supervisory concern in one or more of the component areas. These financial institutions exhibit a combination of weaknesses that may range from moderate to severe; however, the magnitude of the deficiencies generally will not cause a component to be rated more severely than 4. Management may lack the ability or willingness to effectively address weaknesses within appropriate time frames. Financial institutions in this group generally are less capable of withstanding business fluctuations and are more vulnerable to outside influences than those institutions rated a composite 1 or 2. Additionally, these financial institutions may be in significant noncompliance with laws and regulations. Risk management practices may be less than satisfactory relative to the institution's size, complexity, and risk profile. These financial institutions require more than normal supervision, which may include formal or informal enforcement actions. Failure appears unlikely, however, given the overall strength and financial capacity of these institutions.

- Composite 4: Financial institutions in this group generally exhibit unsafe and unsound practices or conditions. There are serious financial or managerial deficiencies that result in unsatisfactory performance. The problems range from severe to critically deficient. The weaknesses and problems are not being satisfactorily addressed or resolved by the board of directors and management. Financial institutions in this group generally are not capable of withstanding business fluctuations. There may be significant noncompliance with laws and regulations. Risk management practices are generally unacceptable relative to the institution's size, complexity, and risk profile. Close supervisory attention is required, which means, in most cases, formal enforcement action is necessary to address the problems. Institutions in this group pose a risk to the deposit insurance fund. Failure is a distinct possibility if the problems and weaknesses are not satisfactorily addressed and resolved.

- Composite 5: Financial institutions in this group exhibit extremely unsafe and unsound practices or conditions; exhibit a critically deficient performance; often contain inadequate risk management practices relative to the institution's size, complexity, and risk profile; and are of the greatest supervisory concern. The volume and severity of problems are beyond management's ability or willingness to control or correct. Immediate outside financial or other assistance is needed in order for the financial institution to be viable. Ongoing supervisory attention is necessary. Institutions in this group pose a significant risk to the deposit insurance fund and failure is highly probable.

A.2 Forms FFIEC 031, 041, and 051: 2000 through 2023

We use the Federal Financial Institutions Examination Council (FFIEC) Consolidated Reports of Condition and Income ("Call Report"), forms FFIEC 031, 041, and 051. These data provide quarterly information on balance sheets and income statements on a consolidated basis for all commercial banks operating in the United States and regulated by the FRS, the FDIC, and the OCC.

A.3 Call Revision Data

Banks submit their Call Report data quarterly through the FFIEC's Central Data Repository (CDR). Their submission must be processed (and accepted) at most 30 calendar days after the end of each quarter.[1] In order for a submission to be accepted by the CDR, it must pass a series of validity and quality checks, known as the "FFIEC Edit Criteria", which help to ensure the report is internally consistent and accounting identities hold. Only once a submission passes the edit criteria and is accepted it will enter the FFIEC's internal system. See the inter-agency "Guidelines for Resolving Edits" for more information.

FFIEC 031 and 041 instructions:

> A bank's primary federal bank supervisory authority may require the filing of an amended Call Report if reports as previously submitted contain significant errors, as determined by the supervisory authority, in how the reporting bank classified or categorized items in the reports, i.e., on what line of the report an item has been reported.
>
> When dealing with the recognition and measurement of events and transactions in the Call Report, amended reports may be required if a bank's primary federal bank supervisory authority determines that the reports as previously submitted contain errors that are material for the reporting bank. Materiality is a qualitative characteristic of accounting information that is addressed in FASB Concepts Statement No. 8, "Conceptual Framework for Financial Reporting," as follows: "Information is material if omitting it or misstating it could influence decisions that users make on the basis of the financial information of a specific reporting entity. In other words, materiality is an entity-specific aspect of relevance based on the nature or magnitude or both of the items to which the information relates in the context of an individual entity's financial report."

Data corrections are allowed after a submission is accepted, including for prior periods. This can occur for multiple reasons, such as if a new call report flags inconsistencies across time in reported amounts.

Table A.1 shows the probability of Call Report revisions from 2007 through 2024 for selected line items.

[1] Banks with multiple non-shell foreign offices are allowed a longer period of 35 calendar days.

Table A.1: *Summary Statistics: Call Report Data Revisions, 2007-2024.*

	Probability of revision	Days to first revision		Numbers of revisions	
		Mean	Median	Mean	Median
Report:					
Entire call report	0.296	114	64	1.4	1.0
Balance sheet	0.266	116	65	1.4	1.0
Income statement	0.123	116	90	1.2	1.0
Line items:					
Total assets	0.052	122	87	1.1	1.0
Non-performing loans	0.018	106	75	1.1	1.0
Loan-loss provisions	0.019	112	88	1.1	1.0
Total equity capital	0.050	124	89	1.1	1.0
Net income	0.046	120	87	1.1	1.0

Notes: This table reports summary statistics on FFIEC call report data corrections between 2007Q2 (the first quarter where this information is broadly available) and 2024Q3. The "Days to first revision" and "Number of revisions" columns are conditional on at least one revision.

B Additional Tables

Table B.2: *Summary Statistics: Bank-level data from 2000 through 2023.*

	N	Mean	Std. dev.	1st	10th	25th	Median	75th	90th	99th
Assets (in million)	637,180	1,559.20	26,128.70	13.89	48.98	93.91	199.27	466.23	1,169.53	15,681.90
Number of Employees	637,179	202.36	2,942.31	4.00	10.00	19.00	38.00	84.00	196.00	1,775.00
Age	637,182	36.73	18.83	1.00	7.00	19.00	43.00	51.00	58.00	63.00
CET 1/RWA	574,581	0.22	2.02	0.07	0.10	0.12	0.14	0.19	0.27	1.54
Tier 1 Capital/RWA	609,338	0.22	1.76	0.08	0.10	0.12	0.14	0.19	0.27	1.50
Equity/RWA	609,338	0.23	1.76	0.09	0.11	0.13	0.15	0.20	0.28	1.50
TCE/assets	634,939	0.11	0.09	0.05	0.07	0.08	0.10	0.12	0.15	0.41
Equity/assets	637,180	0.12	0.16	0.06	0.08	0.09	0.10	0.12	0.15	0.74
Net income/assets	637,175	0.01	0.02	-0.02	0.00	0.00	0.00	0.01	0.01	0.03
Liquid assets/assets	636,842	0.30	0.17	0.03	0.11	0.18	0.27	0.39	0.53	0.83
Deposits/loans	637,176	0.82	0.12	0.01	0.73	0.80	0.85	0.88	0.90	0.93
Time deposits/assets	637,176	0.33	0.15	0.00	0.13	0.22	0.34	0.44	0.53	0.69
Compensation/empl.	635,665	40,728.02	32,238.92	8,232.14	13,219.51	20,833.33	35,357.14	53,141.30	73,316.24	130771.43

Notes: This table reports summary statistics for the bank-level data based on the FFIEC Call Report.

Table B.3: *Summary Statistics: Bank-level data from 2000 through 2023 for failing banks.*

	N	Mean	Std. dev.	1st	10th	25th	Median	75th	90th	99th
Assets (in million)	18,310	3,041.76	34,190.30	16.36	49.46	109.14	240.52	554.55	1,508.55	52,833.29
Number of Employees	18,310	488.64	7,032.51	5.00	12.00	21.00	42.00	94.00	226.00	2,096.00
Age	18,310	22.36	17.37	1.00	3.00	6.00	17.00	41.00	47.00	54.00
CET 1/RWA	16,902	0.12	0.11	0.01	0.06	0.09	0.10	0.13	0.18	0.43
Tier 1 Capital/RWA	18,310	0.12	0.11	0.01	0.07	0.09	0.11	0.13	0.18	0.45
Equity/RWA	18,310	0.14	0.11	0.03	0.08	0.10	0.12	0.14	0.20	0.46
TCE/assets	18,310	0.09	0.05	0.01	0.05	0.07	0.08	0.10	0.13	0.29
Equity/assets	18,310	0.09	0.06	0.01	0.05	0.07	0.09	0.10	0.13	0.31
Net income/assets	18,310	-0.00	0.02	-0.07	-0.02	-0.00	0.00	0.01	0.01	0.02
Liquid assets/assets	18,310	0.19	0.12	0.02	0.06	0.11	0.17	0.25	0.37	0.61
Deposits/loans	18,310	0.83	0.11	0.40	0.72	0.80	0.85	0.89	0.92	0.97
Time deposits/assets	18,310	0.47	0.16	0.03	0.26	0.36	0.47	0.58	0.67	0.82
Compensation/empl.	18,301	38,017.47	24,779.59	7,689.19	13,052.17	19,783.78	33,888.89	50,035.71	66,214.29	116308.52

Notes: This table reports summary statistics for the bank-level data based the FFIEC Call Report.

Table B.4: *Balance Test: Bank Characteristics by Expected Exam Lead Early in GFC*

	Expect State		Expect Federal		Difference	
	Mean	Std	Mean	Std	Diff	t-stat
CAMELS Rating	1.575	0.495	1.581	0.494	0.006	0.346
Size (Assets, in log)	18.705	1.227	18.762	1.147	0.057	1.440
Equity/Assets	0.111	0.058	0.110	0.051	-0.002	-1.023
Net income/Assets	0.010	0.015	0.009	0.013	-0.000	-1.070
Deposits/Assets	0.817	0.091	0.814	0.089	-0.003	-1.058
Time Deposits/Assets	0.393	0.123	0.390	0.115	-0.003	-0.873
Wholesale Funding/Assets	0.046	0.058	0.054	0.098	0.008	0.862
Loans/Assets	0.655	0.163	0.662	0.160	0.007	1.248
NPL/Loans	0.012	0.013	0.012	0.012	0.000	0.501

This table reports selected observable characteristics as of 2007Q4 across bank that were expected to receive either a state or a federal-led exam early in the GFC. Standard errors are clustered at the bank level.

Table B.5: *Effect of Supervisory Strictness on CAMELS and the Number and Length of On-site Examinations during the GFC: OLS*

Dependent variable	CAMELS	Prob of CAMELS≥3	No. of FRS/FDIC exam	No. of exams	Length of exams
	(1)	(2)	(3)	(4)	(5)
High supervisory strictness	0.24***	0.090***	0.84***	0.54***	3.26***
	(0.023)	(0.0094)	(0.059)	(0.064)	(0.96)
Observations	3604	3604	3604	3604	3575
Mean dep. var	2.05	0.21	2.88	4.59	49.0
R^2	0.37	0.29	0.29	0.23	0.31
State FE	✓	✓	✓	✓	✓
Baseline Controls	✓	✓	✓	✓	✓

Notes: This table reports results from estimating (3). The outcome, y_b, is either the bank's average CAMELS rating, the probability of being rated with a 3 or higher, the number of total exams, the number of exams led by a federal agency (FDIC or FRS), or the length of the average exam (in days) from 2008q3 through 2013q4. High supervisory strictness$_b$ is a dummy that indicates whether bank b was subject to at least one exam led by a federal agency in the early phase of the GFC, defined as the 18 months following the failure of Lehman Brothers. We provide both OLS and 2SLS estimates. The sample is restricted to state-member banks and non-member banks that are supervised by either the Federal Reserve, the FDIC, or state regulatory agencies as well as to banks that qualified for the rotating schedule before 2008q3. Baseline controls are a bank's size, equity over assets, time deposits over assets, non-performing loans over total loans, and net income over assets as of 2008q2. Robust standard errors. *,**, and *** indicate significance at the 10%, 5%, and 1% level, respectively.

Table B.6: ***Loss Rates for Uninsured Depositors in Bank Failure***

Era	Number of failures	Share of failures with losses to depositors	Conditional loss rate	Unconditional loss rate
1992-2007	302	0.43	0.24	0.10
2008-2022	536	0.06	0.43	0.03
All	838	0.2	0.28	0.06

Notes: The data are as reported in FDIC (2023). The conditional loss rate is the loss rate for failures involving a loss for uninsured depositors.

Table B.7: *Costs for the FDIC Insurance Fund and Speed of Failure*

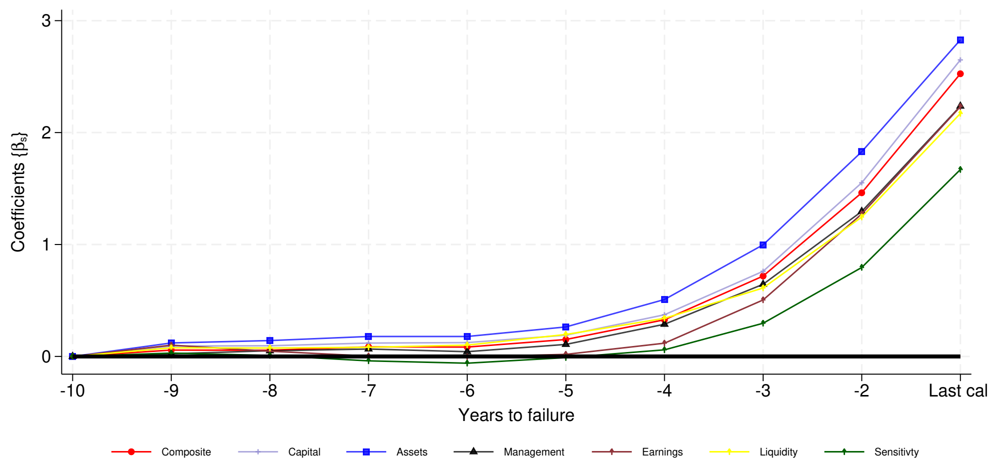
Dependent variable	DIF costs/assets			
	OLS	2SLS	OLS	2SLS
	(1)	(2)	(3)	(4)
High supervisory strictness	-0.067**	-0.12**	-0.080**	-0.13**
	(0.031)	(0.052)	(0.032)	(0.053)
Speed of failure	-0.000065**	-0.000073**	-0.000037	-0.000031
	(0.000028)	(0.000028)	(0.00011)	(0.00011)
Observations	116	116	116	116
Mean dep. var	0.22	0.22	0.22	0.22
R^2	0.54	0.19	0.56	0.20
State FE	✓	✓	✓	✓
Baseline Controls	✓	✓	✓	✓
Year of failure FE			✓	✓
First-stage F-stat		16.1		15.7

Notes: This table reports results from estimating (3), where y_b is the cost the FDIC incurred as a share of the failed banks assets at failure. High supervisory strictness$_b$ is a dummy that indicates whether bank b was subject to at least one exam led by a federal agency in the early phase of the GFC, defined as the 18 months following the failure of Lehman Brothers. Speed of failure$_b$ is the number of days elapsed since between the failure of Lehman Brothers and bank b. We provide both OLS and 2SLS estimates. In the 2SLS model, we instrument whether a bank is subject to a federal exam in the early phase of the GFC with whether it was having rotating exams in the four previous exams. The sample is restricted to state-member banks and non-member banks that are supervised by either the Federal Reserve, the FDIC, or state regulatory agencies as well as to banks that qualified for the rotating schedule before September 2008. Baseline controls are a bank's size, equity over assets, time deposits over assets, non-performing loans over total loans, and net income over assets as of 2008q2. Robust standard errors. *,**, and *** indicate significance at the 10%, 5%, and 1% level, respectively.

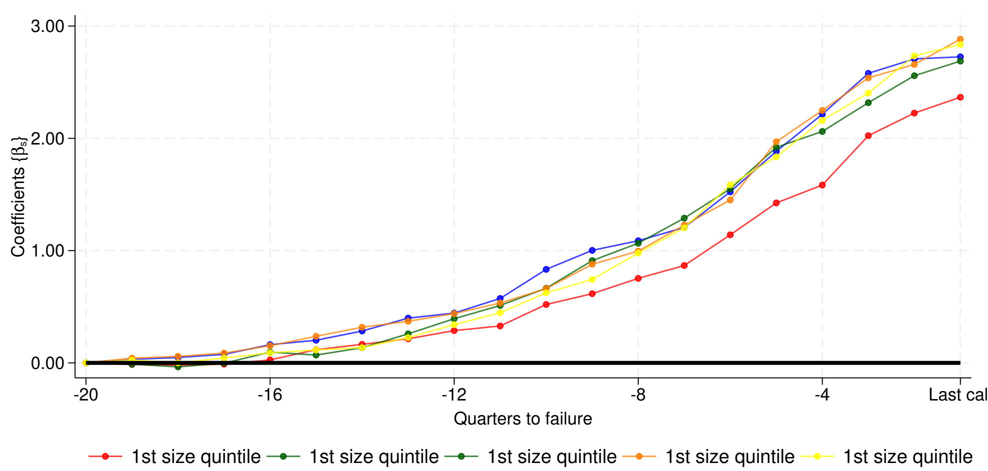
C Additional Figures

Figure C.1: ***CAMELS Ratings by Type and Bank Size in Failing Banks: 2000-2023***

(a) *CAMELS in failing banks by rating type*

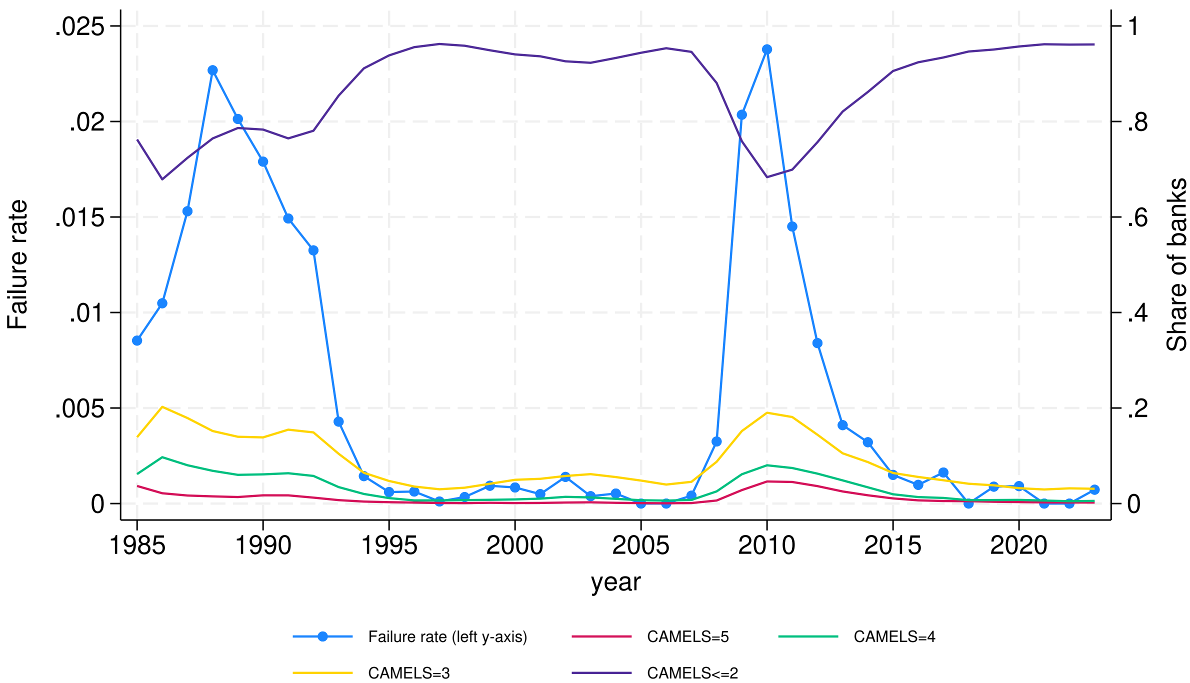


(b) *CAMELS in failing banks by bank size*



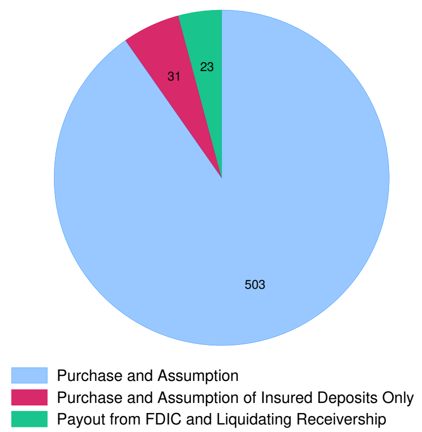
Notes: The figure presents the sequence of coefficients from estimating Equation (1), where the dependent variable is the ratio indicated in the figure legend. The specification includes a set of bank fixed effects. The sample is restricted to failing banks that fail sometime between 2000 and 2023 and includes data for the 5 years before failure.

Figure C.2: ***CAMELS Ratings and Bank Failure Rates***



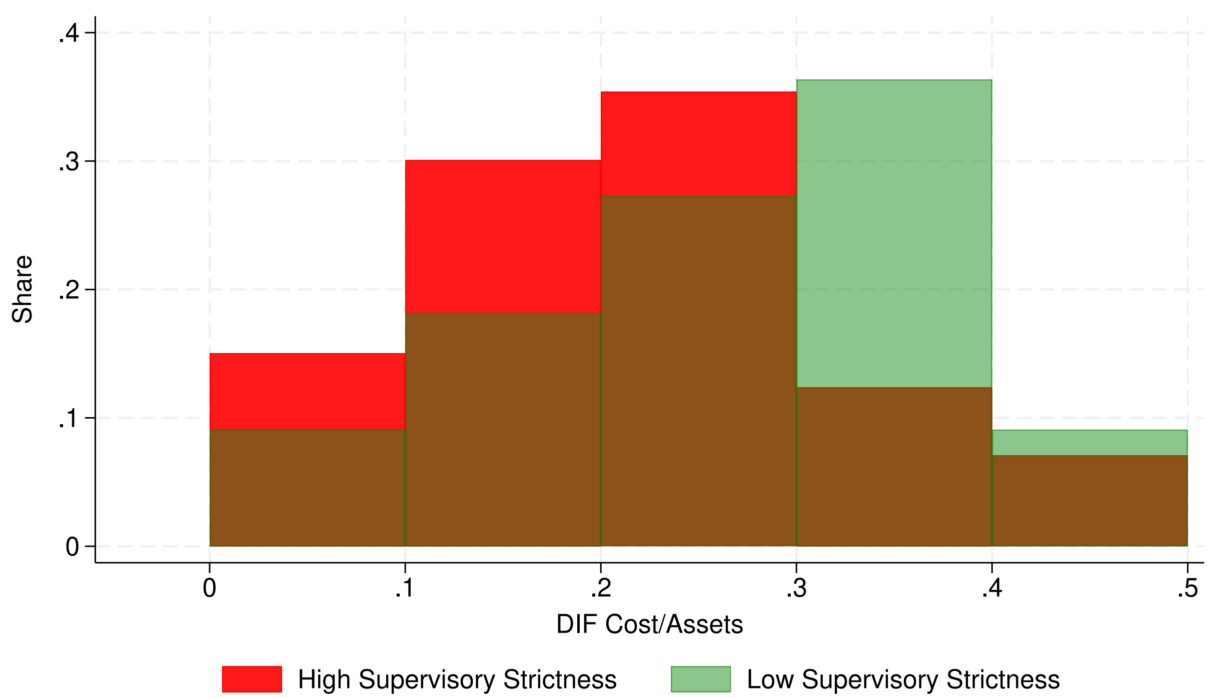
Notes: This figure shows the distribution of CAMELS ratings (right y-axis) and the realized rate of bank failures (left y-axis) from 1985 through 2023.

Figure C.3: ***Resolution of Failing Banks by Resolution Type, 2000-2023***



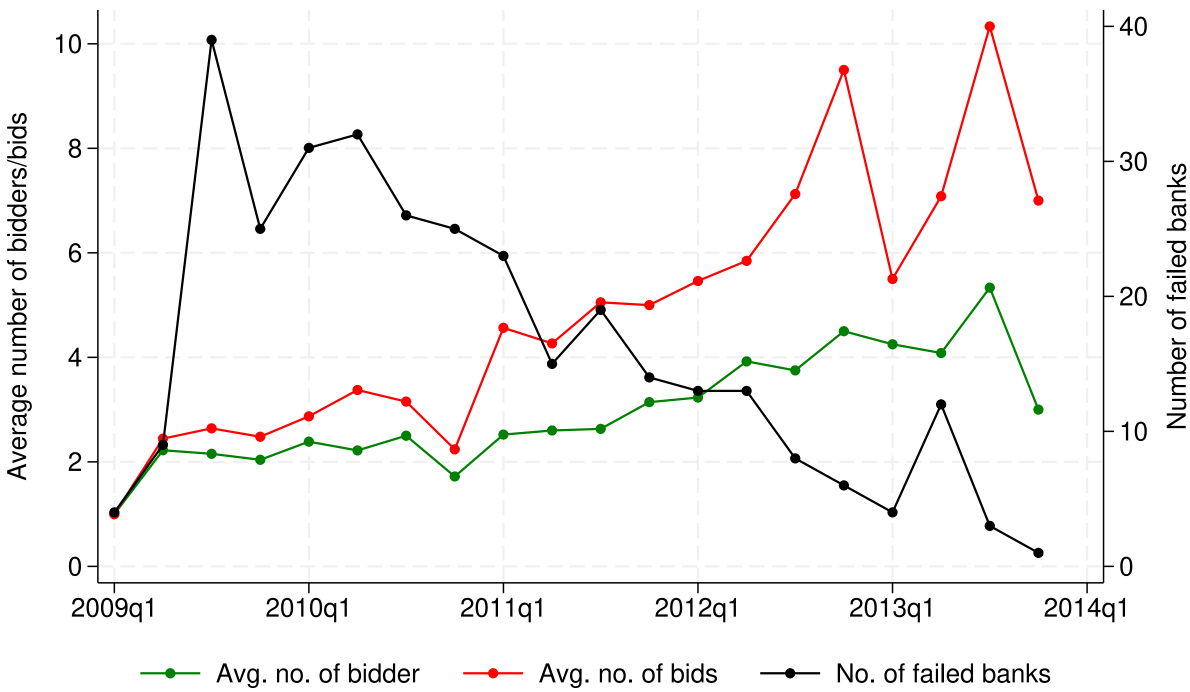
Notes: This figure shows the count of resolution types of FDIC member banks that failed between 2000 and 2023. There are three types of failures. PA: "Purchase and Assumption", where the insured and uninsured deposits, certain other liabilities and a portion of the assets were sold to an acquirer. PI: "Purchase and Assumption of the insured deposits only", where the traditional P&A was modified so that only the insured deposits were assumed by the acquiring institution. And, PO: "Payout", where the insurer paid the depositors directly and placed the assets in a liquidating receivership.

Figure C.4: ***Cost of Failure to the FDIC DIF by Supervisory Strictness***



Notes: This figure plots a histogram of the cost of failure to the FDIC DIF for failures during the GFC (measured between 2008q3-2013q4) by whether banks were exogenously exposed to high or low supervisory strictness as the onset of the GFC. The sample is restricted to state-member banks and non-member banks that are supervised by either the Federal Reserve, the FDIC, or state regulatory agencies as well as to banks that qualified for the rotating schedule before September 2008.

Figure C.5: ***Number of Bidders on Failed Banks during the GFC***



Notes: This figure plots the number of failed banks, the average number of bidders, and the average number of bids in FDIC failed bank auctions. The data are from Allen et al. (2023a).

D Additional Results

D.1 The Predictive Power of Supervisory Rating for Bank Failure

In this section, we study the predictive power of CAMELS ratings more formally. We consider the predictive power at different predictions horizons and compare it to the predictive power of financial information that can be obtained in publicly available call report.

To more formally understand whether CAMELS captures bank failure risk, we estimate simple predictive regression models of the following form:

$$\text{Failure}_{b,t+1} = \alpha + \sum_{k=1}^{5} \beta_k \times \mathbf{1}[\text{CAMELS}_{b,t} = k] + \Gamma X_{b,t} + \epsilon_{b,t+1}, \tag{A.1}$$

where $\text{Failure}_{b,t+1}$ is an indicator variable equal to one if bank b fails over the next year, and $\mathbf{1}[\text{CAMELS}_{bt} = k]$ is an indicator variable that equals one when bank b has a CAMELS rating of k at time t, with k ranging from 1 through 5. In addition to CAMELS ratings, we include information from banks' publicly available financial statements, $X_{b,t}$. This allows us to compare the predictive power of supervisory ratings with the predictive power of information from financial statements. Specifically, we include income-to-assets to proxy for a bank's risk of insolvency, time deposits-to-assets to proxy for a bank's reliance on relatively expensive noncore funding, and quintiles of bank loan growth to proxy bank credit growth and risk-taking.[2] Finally, we also consider the interaction between the insolvency measure and noncore funding, as well as interactions of all of the variables from financial statements with the CAMELS rating.

To quantify the power of these observables for predicting bank failure, we construct the receiver operating characteristic (ROC) curve. The ROC curve is a standard tool used to evaluate binary classification ability. The ROC curve traces out the true positive rate against the false positive rate as we vary the classification threshold. We then calculate the area under the ROC curve (AUC). An uninformative predictor has an AUC of 0.5, whereas an informative predictor has an AUC of greater than 0.5. The AUC metric is commonly used in the literature on predicting financial crises.[3] Furthermore, we test both in-sample and pseudo-out-of-sample classification performance. The pseudo-out-of-sample AUC is constructed by estimating Equation (A.1) iteratively on an expanding sample and predicting the probability of failure for each bank in the h years after the call report measured at time t, using only data up to year t. In addition to the AUC, in Appendix Table D.12 we also provide the precision-recall curve, an alternative metric

[2]These measures have been shown to predict failure with a high degree of accuracy (Correia et al., 2024). We calculate quintiles of loan growth using the distribution of loan growth up to year t to avoid look-ahead bias.

[3]For reference, the in-sample AUC for predicting financial crises in aggregate data based on credit and asset price growth is typically in the 0.65–0.75 range (see e.g., Schularick and Taylor, 2012; Drehmann and Juselius, 2014; Baron et al., 2021; Greenwood et al., 2022; Müller and Verner, 2024). Correia et al. (2024) further show that the AUC when predicting individual bank failures is around 0.86 for the historical, pre-FDIC U.S. banking system and 0.95 for the contemporary U.S. banking system.

that is well-suited to evaluating classification performance for rare outcomes. The basic findings are robust to this alternative metric.

Before proceeding, we make two remarks on the interpretation of these predictive regressions. First, CAMELS ratings may not anticipate all failures if supervisors can act to avert some failures, for example, by requiring troubled banks to raise equity. Second, we note that the comparison between the predictive performance of CAMELS ratings and metrics from financial statements must be interpreted within the context of a specific equilibrium. In particular, as we show below, supervisors play an important role in auditing and ensuring the accuracy of bank financial statements. Thus, a high predictive performance of financial metrics does not necessarily imply that a similarly strong performance could be obtained in a regime with less supervisory scrutiny, as the quality of financial statements could deteriorate.

Table D.8 shows that CAMELS bank supervisory ratings are highly predictive of bank failure at short horizons. The AUC is 96% at the one-year horizon (see column 1). Figure D.8 plots the ROC curve for predicting failure at the one-year horizon—corresponding to Panel A of Table D.8. A model based on confidential CAMELS ratings can identify around 75% of true positives with less than 1% false positives and close to 90% of true positives with around 5% false positives. This evidence indicates that supervisors typically do anticipate bank failures in this sample.

Previous work has established that bank failures in the contemporary U.S. banking system are substantially predictable based on variables capturing income, capitalization, funding structure, and credit growth (see, e.g., Cole and Gunther, 1998; Wheelock and Wilson, 2000; Cole and White, 2012; Correia et al., 2024). At the one-year horizon, we find that the AUC (both in- and out-of-sample) is slightly higher when using a bank's confidential supervisory rating as the predictor, compared to when using measures of bank solvency and funding vulnerabilities (see columns 2-4). Moreover, interacting measures of bank solvency and funding and the CAMELS rating can improve the AUC further, especially at longer prediction horizons (column 5). These findings suggest that supervisory ratings contain some soft information for predicting failure in the near term that is not yet reflected in financial statements, in addition to capturing the information in public financial statements.[4]

We also consider the ability of CAMELS to predict failures at longer horizons. Panel B in Table D.8 shows that when predicting whether a bank will fail in two to three years, an econometric specification based on bank financial metrics slightly outperforms CAMELS ratings. Moreover, Panel C shows that when predicting failure in four to five years, an econometric specification based solely on bank credit growth outperforms the CAMELS ratings, both in- and out-of-sample. This is consistent with credit growth being a strong leading indicator of bank failures, see Figure D.6, and banking crises (e.g. Schularick and Taylor, 2012; Greenwood et al., 2022; Müller and Verner, 2024). The outperformance of

[4]Note that when we merge CAMELS ratings and Call Report data, we carry ratings forward until an examination leads to a change in a CAMELS rating. This gives rise to the potential concern that because financial data are updated quarterly and supervisory ratings are updated only every 12 months or 18 months, the informational value of CAMELS ratings may be lower outside of examination dates (see, e.g., Cole and Gunther, 1998; Hirtle and Lopez, 1999). In additional unreported regressions, we restrict the sample to examination dates only and find similar levels and differences in the AUC across specifications.

Table D.8: *Predicting Bank Failures with Supervisory Ratings and Bank Fundamentals*

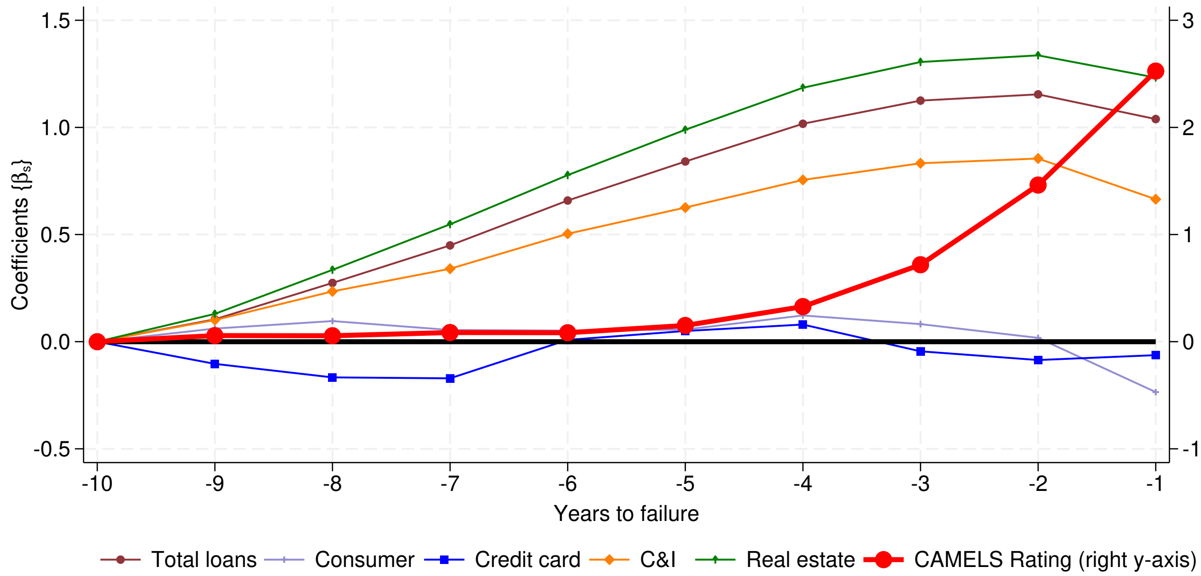
	(1)	(2)	(3)	(4)	(5)	(6)
Panel A: Failure in $t+1$						
AUC (in-sample)	0.961	0.935	0.767	0.936	0.963	0.936
AUC (out-of-sample)	0.960	0.933	0.722	0.936	0.957	0.929
N	633759	633751	585720	585717	585717	630437
No of Banks	11222	11222	10749	10749	10749	11221
Mean of dep. var.	.29	.29	.29	.29	.29	.29
Panel B: Failure between $t+2$ and $t+3$						
AUC (in-sample)	0.835	0.845	0.743	0.867	0.897	0.844
AUC (out-of-sample)	0.762	0.808	0.671	0.832	0.858	0.801
N	633759	633751	585720	585717	585717	630437
No of Banks	11222	11222	10749	10749	10749	11221
Mean of dep. var.	.59	.59	.55	.55	.55	.58
Panel C: Failure between $t+4$ and $t+5$						
AUC (in-sample)	0.750	0.771	0.781	0.816	0.837	0.770
AUC (out-of-sample)	0.656	0.723	0.738	0.778	0.790	0.724
N	633759	633751	585720	585717	585717	630437
No of Banks	11222	11222	10749	10749	10749	11221
Mean of dep. var.	.55	.55	.5	.5	.5	.55
Specification details						
CAMELS	✓				✓	
Insolvency × Non-core funding		✓		✓	✓	✓
Credit growth			✓	✓	✓	
Insolvency × Non-core funding × CAMELS					✓	
Age controls	✓	✓	✓	✓	✓	✓
Excluding data revisions						✓

Notes: This table reports the area under the receiver operating characteristic curve (AUC) across different specifications and horizons using in-sample and pseudo-out-of-sample classification. The regression coefficients underlying the estimation in columns (1) through (4) are reported in Table D.9 (Panel A), Table D.10 (for Panel B), and Table D.11 (for Panel C) in the Appendix. See Table D.12 below for precision-recall statistics.

financial metrics is even stronger when using a range of balance sheet metrics (column 4 in Panel C). The difference is substantial, with an out-of-sample AUC of 77% for the model based on financial metrics, compared to 66% for the model based on CAMELS. Overall, the evidence indicates that supervisory CAMELS ratings have strong predictability for failures in the near term, but other financial metrics, especially credit growth, contain additional information at longer horizons.

The worse performance of CAMELS relative to credit growth at longer horizons suggests CAMELS does not fully incorporate the medium-run risks associated with rapid

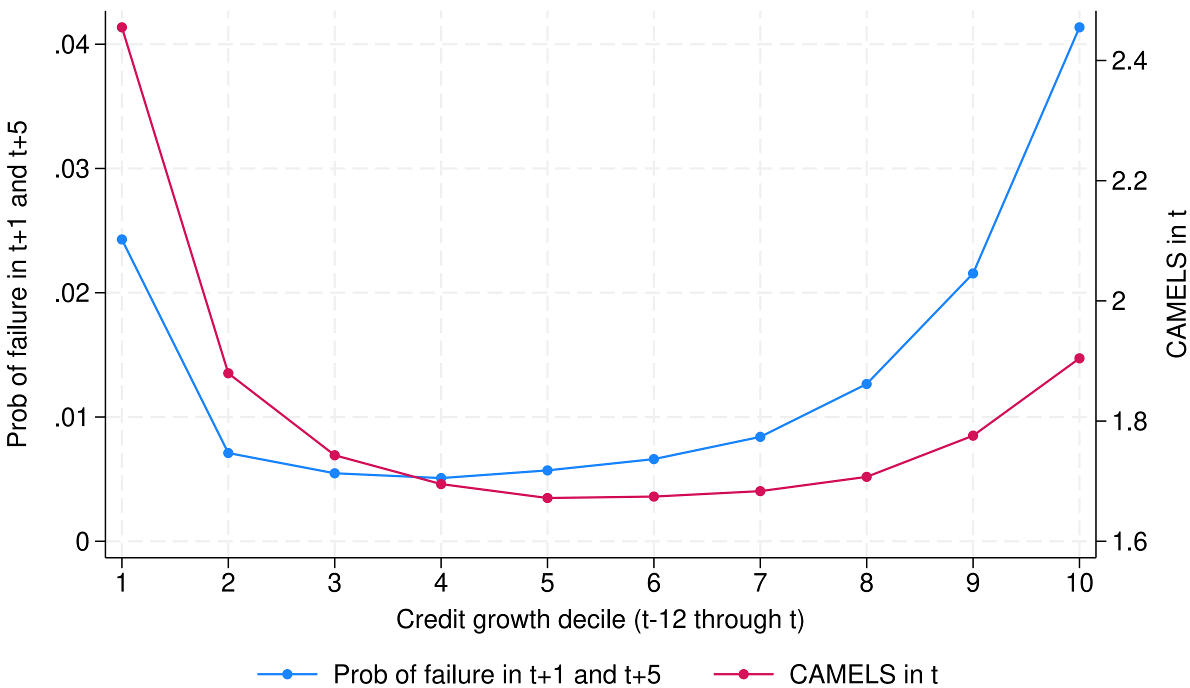
Figure D.6: ***Loan Growth in Failing Banks: 2000-2023***



Notes: The figure presents the sequence of coefficients from estimating Equation (1), where the dependent variable is the log of the type of loan indicated in the figure legend. The specification includes a set of bank fixed effects. The sample is restricted to the 524 failing banks that fail between 2000 and 2023 and includes data for the 5 years before failure.

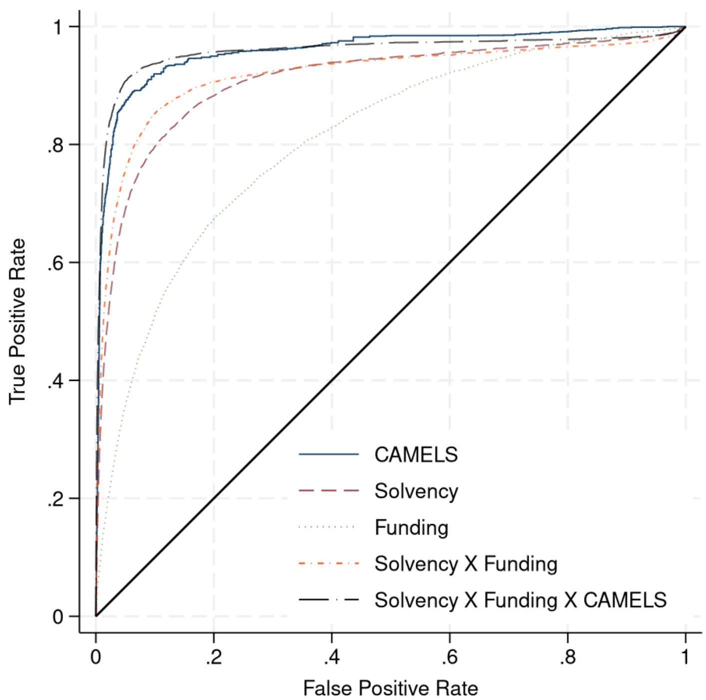
credit growth. In line with this intuition, Figure D.7 shows that, while the probability of failure over the next five years is highest in banks with the fastest loan growth over the past three years, CAMELS ratings only rise slightly across loan growth quintiles. Moreover, Figure C.2 shows that the share of troubled banks with a rating of 3 to 5 remained historically low between 2000 and 2007, during a period of historically rapid credit growth. This evidence indicates that CAMELS ratings have historically not reflected the risk of rapid loan growth for medium-term failure. This may be because supervisors neglect downside risks, much like bank investors (Baron and Xiong, 2017; Fahlenbrach et al., 2018). At the same time, it may also reflect incentive problems. Supervisors may face the challenge that type I errors based on loan growth (i.e., restricting strong banks with sound businesses from growing) are particularly economically and politically costly.

Figure D.7: ***Three-Year Credit Growth, Probability of Failure, and CAMELS***



Notes: This figure plots the probability of failure over the next five years (left y-axis) and CAMELS this year (right y-axis) across deciles of loan growth over the past 12 quarters (three-years). Data are from 2000 onwards.

Figure D.8: ***ROC Curves for Predicting Bank Failure: 2000-2023***



Notes: This figure plots the receiver operating characteristic (ROC) curve for the estimates based on columns (1) through (4) in Table D.8 Panel C.

Table D.9: ***Predicting Bank Failures: 2000-2023***

Horizon h	Fail within next year			
	(1)	(2)	(3)	(4)
Regulatory Rating:				
- CAMELS = 2	0.00			
	(0.00)			
- CAMELS = 3	0.21**			
	(0.00)			
- CAMELS = 4	2.02***			
	(0.01)			
- CAMELS = 5	19.26***			
	(0.03)			
Solvency:				
- Net Income/Assets		9.23***		35.16***
		(0.02)		(0.07)
Funding:				
- Time Deposits/Deposits		4.04***		4.73***
		(0.01)		(0.01)
- NI/Assets $\times$ TD/Dep.		-222.29***		-289.79***
		(0.31)		(0.42)
Credit Growth:				
- Q1 of credit growth from t-3 to t			0.57***	0.18***
			(0.00)	(0.00)
- Q2 of credit growth from t-3 to t			0.01	-0.04***
			(0.00)	(0.00)
- Q4 of credit growth from t-3 to t			-0.01	0.01
			(0.00)	(0.00)
- Q5 of credit growth from t-3 to t			-0.05	-0.02
			(0.00)	(0.00)
N	633759	633751	585720	585717
No of Banks	11222	11222	10749	10749
Mean of dep. var.	.29	.29	.29	.29

Notes: This table presents OLS estimates of (A.1) with failure within the next year as the dependent variables for the 2000-2023 sample. In addition to the reported predictor variables, we also include the log of a bank's age. Standard errors in parentheses are clustered at the bank level; *, **, and *** indicate significance at the 10%, 5%, and 1% level, respectively.

Table D.10: ***Predicting Bank Failures: 2000-2023***

Horizon h	Fail within next three years			
	(1)	(2)	(3)	(4)
Regulatory Rating:				
- CAMELS = 2	0.23**			
	(0.00)			
- CAMELS = 3	1.11***			
	(0.00)			
- CAMELS = 4	4.42***			
	(0.01)			
- CAMELS = 5	10.78***			
	(0.01)			
Solvency:				
- Net Income/Assets		5.11***		15.25***
		(0.01)		(0.02)
Funding:		ref.		ref.
- Time Deposits/Deposits		5.05***		5.07***
		(0.01)		(0.01)
- NI/Assets $\times$ TD/Dep.		-146.55***		-174.39***
		(0.18)		(0.19)
Credit Growth:				
- Q1 of credit growth from t-3 to t			0.27***	0.01
			(0.00)	(0.00)
- Q2 of credit growth from t-3 to t			-0.02	-0.06***
			(0.00)	(0.00)
- Q4 of credit growth from t-3 to t			0.08*	0.09**
			(0.00)	(0.00)
- Q5 of credit growth from t-3 to t			0.72***	0.73***
			(0.00)	(0.00)
N	633759	633751	585720	585717
No of Banks	11222	11222	10749	10749
Mean of dep. var.	.59	.59	.55	.55

Notes: This table presents OLS estimates of (A.1) with failure between $t+1$ and $t+3$ as the dependent variables for the 2000-2023 sample. In addition to the reported predictor variables, we also include the log of a bank's age. Standard errors in parentheses are clustered at the bank level; *, **, and *** indicate significance at the 10%, 5%, and 1% level, respectively.

Table D.11: ***Predicting Bank Failures: 2000-2023***

Horizon h	Fail within next five years			
	(1)	(2)	(3)	(4)
Regulatory Rating:				
- CAMELS = 2	0.36***			
	(0.00)			
- CAMELS = 3	0.49***			
	(0.00)			
- CAMELS = 4	1.12***			
	(0.00)			
- CAMELS = 5	1.50***			
	(0.00)			
Solvency:				
- Net Income/Assets		0.27		1.04
		(0.00)		(0.01)
Funding:				
- Time Deposits/Deposits		2.73***		2.38***
		(0.01)		(0.00)
- NI/Assets × TD/Dep.		-8.11		-11.15*
		(0.07)		(0.07)
Credit Growth:				
- Q1 of credit growth from t-3 to t			0.01	-0.01
			(0.00)	(0.00)
- Q2 of credit growth from t-3 to t			-0.04*	-0.04**
			(0.00)	(0.00)
- Q4 of credit growth from t-3 to t			0.22***	0.21***
			(0.00)	(0.00)
- Q5 of credit growth from t-3 to t			1.17***	1.16***
			(0.00)	(0.00)
N	633759	633751	585720	585717
No of Banks	11222	11222	10749	10749
Mean of dep. var.	.55	.55	.5	.5

Notes: This table presents OLS estimates of (A.1) with failure between $t+1$ and $t+5$ as the dependent variables for the 2000-2023 sample. In addition to the reported predictor variables, we also include the log of a bank's age. Standard errors in parentheses are clustered at the bank level; *, **, and *** indicate significance at the 10%, 5%, and 1% level, respectively.

Table D.12: *Predicting Bank Failures with Supervisory Ratings and Bank Fundamentals: Precision Recall*

	(1)	(2)	(3)	(4)	(5)	(6)
Panel A: Prediction horizon $t+1$						
PR-AUC (in-sample)	0.171	0.180	0.012	0.189	0.348	0.162
PR-AUC (out-of-sample)	0.190	0.171	0.009	0.176	0.324	0.155
Mean of dep. var.	0.003	0.003	0.003	0.003	0.003	0.003
Ratio (in-sample)	58.902	61.913	4.006	64.889	119.404	56.667
Ratio (out-of-sample)	50.678	45.739	2.434	46.888	86.383	41.911
Precision at 10% recall (in-smp.)	0.220	0.349	0.021	0.373	0.562	0.320
Precision at 10% recall (o.o.s.)	0.276	0.292	0.013	0.295	0.503	0.280
Panel B: Prediction horizon $t+2 \to t+3$						
PR-AUC (in-sample)	0.057	0.068	0.015	0.073	0.086	0.057
PR-AUC (out-of-sample)	0.051	0.055	0.012	0.058	0.071	0.045
Mean of dep. var.	0.006	0.006	0.006	0.006	0.006	0.006
Ratio (in-sample)	9.627	11.535	2.661	13.233	15.555	9.716
Ratio (out-of-sample)	6.557	7.166	1.732	8.111	9.880	5.914
Precision at 10% recall (in-smp.)	0.114	0.142	0.020	0.142	0.166	0.108
Precision at 10% recall (o.o.s.)	0.084	0.107	0.015	0.115	0.125	0.086
Panel C: Prediction horizon $t+4 \to t+5$						
PR-AUC (in-sample)	0.015	0.020	0.017	0.025	0.027	0.020
PR-AUC (out-of-sample)	0.012	0.015	0.014	0.019	0.019	0.015
Mean of dep. var.	0.006	0.006	0.005	0.005	0.005	0.006
Ratio (in-sample)	2.629	3.631	3.438	5.011	5.352	3.656
Ratio (out-of-sample)	1.803	2.299	2.388	3.098	3.195	2.301
Precision at 10% recall (in-smp.)	0.020	0.037	0.022	0.042	0.046	0.039
Precision at 10% recall (o.o.s.)	0.018	0.020	0.017	0.022	0.024	0.020
Specification details						
CAMELS	✓				✓	
Insolvency × Non-core funding		✓		✓	✓	✓
Credit growth			✓	✓	✓	
Insolvency × Non-core funding × CAMELS					✓	
Age controls	✓	✓	✓	✓	✓	✓
Excluding data revisions						✓

Notes: This table reports the area under the precision-recall curve (PR-AUC) across different specifications and horizons using in-sample and pseudo-out-of-sample classification. The regression coefficients underlying the estimation of column (1) through (4) are reported in Table D.11 (for Panel A), Table D.10 (Panel B), and Table D.9 (Panel C) in the Appendix.

D.2 Capitalization and Enforcement Action

In this section, we discuss additional insights on the role of banking capitalization and enforcement actions. We first discuss the importance of a bank being undercapitalized in a legal sense for enforcement actions and show how the timing of undercapitalization can explain the absence of enforcement actions in some failing banks. Second, we show how the shaping of financial statements by supervisors matters for making enforcement actions legally permissible.

Timing of Enforcement and Capitalization To illustrate the importance of banking capitalizaton for enforcement actions, we estimate the following regression:

$$\text{Enforcement}_{bt} = \text{Capitalization}_{bt-1} + \epsilon_{bt},$$

where Enforcement_{bt} is a dummy indicating whether bank b was subject to a form of public enforcement at time t, and $\text{Capitalization}_{bt-1}$ is dummy indicating whether bank b is "well capitalized," "adequately capitalized," or "undercapitalized" and, if undercapitalized, whether it is "significantly" or "critically" undercapitalized.[5]

We find that a key determinant of enforcement action is a bank's capitalization. Table D.13 shows results from regressing a dummy whether bank is subject to enforcement action in a given quarter on its capitalization in the previous quarter. We find that the probability of being subject to enforcement action monotonically increases in a bank's capitalization. But this monotonicity masks some underlying heterogeneity. For instance, while the unconditional probability of being subject to PCA in any given quarter is only 0.3% and only modestly higher for in adequately banks compared to well capitalized banks, we find that significantly undercapitalized have a 6 percentage points and critically undercapitalized a 10 percentage point and this a much higher chance of being subject to PCA enforcement. Cease and Desist order or written agreement, in turn, are most common in adequately but not well or undercapitalized banks. The patterns suggest that supervisors first resort to relatively mild forms of enforcement action as a banks capitalization erodes but start to take more drastic action as default becomes a material risk. Figure D.11 below confirms this pattern when considering enforcement actions by CAMELS ratings.

While capitalization is a good candidate to explain whether enforcement action takes place or not, the timing and extend of a bank's decline in capitalization is also important. We find that failing banks tend to cross the legal thresholds that trigger drastic enforcement action such as PCA relatively late and close to the failure date. Figure D.9 plots results from estimating Equation (1) on a bank's capitalization. Figure D.9 reveals that failing banks become undercapitalized to an extent that requires formal legal action by supervisors only in the months right before failure. This short interval leaves little time for enforcement action to be formalized. For instance, Table D.14 shows that, on average, only 20% of all failing banks are critically undercapitalized within one year of failure, and only around 57% are critically undercapitalized within one quarter of failure. Thus, one potential explanation for the absence of enforcement action in some

[5]The detailed definitions can be found here.

Table D.13: *Enforcement actions and bank capitalization.*

Dependent Variable	Combined	PCA	Cease and Desist	Written	Informal
	(1)	(2)	(3)	(4)	(5)
Constant	0.67***	0.0015	0.27***	0.15***	0.22***
	(0.0034)	(0.0012)	(0.0028)	(0.0013)	(0.0010)
Adequately Capitalized	2.62***	0.13***	2.47***	0.25**	-0.15*
	(0.26)	(0.048)	(0.22)	(0.100)	(0.080)
Significantly Undercapitalized	5.09***	6.24***	-0.33	-0.40***	0.074
	(1.89)	(1.61)	(0.89)	(0.099)	(0.43)
Critically Undercapitalized	9.54***	10.0***	1.59	-0.22	-0.40***
	(2.01)	(1.71)	(1.22)	(0.35)	(0.088)
Observations	585553	585553	585553	585553	585553
Mean dep. var	0.74	0.029	0.32	0.16	0.22
R^2	0.045	0.093	0.035	0.032	0.042
Bank FE	✓	✓	✓	✓	✓
Time FE	✓	✓	✓	✓	✓

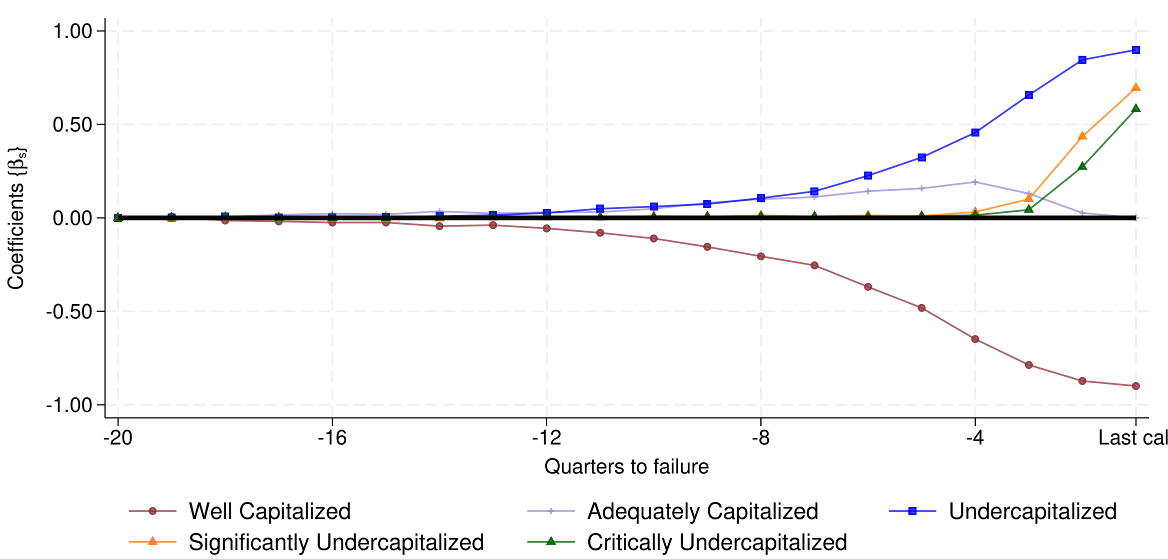
Notes: This table reports results from estimating a model of the following form:

$$\text{Enforcement}_{bt} = \text{Capitalization}_{bt-1} + \epsilon_{bt},$$

where Enforcement_{bt} is a dummy indicating whether bank b was subject to a form of public enforcement at time t, and $\text{Capitalization}_{bt-1}$ is dummy indicating whether bank b is well (coefficient omitted), adequately, under or significantly undercapitalized. Robust standard errors. *,**, and *** indicate significance at the 10%, 5%, and 1% level, respectively.

failing banks, especially the most drastic forms of enforcement, is that failing banks, even as their solvency deteriorates for many years before failure, only become substantially undercapitalized right before failure, leaving little time for formal enforcement action.

Figure D.9: ***Capitalization in Failing Banks***



Notes: This figure plots the set of coefficients $\{\beta_t\}$ from estimating Equation (1) when using a dummy variable indicating whether a bank is well-, adequately-, or undercapitalized; and if undercapitalized whether it's significantly or critically undercapitalized.

Table D.14: *Bank Capitalization Before Failure.*

Well Capitalized	Adequately Capitalized	Undercapitalized		
		Total	Significantly	Critically
A. All banks				
0.98	0.01	0.01	0.00	0.00
B. Failing banks (within 5 years of failure)				
0.70	0.09	0.21	0.06	0.04
C. Failing banks (within 1 year of failure)				
0.17	0.12	0.71	0.29	0.21
D. Failing banks (within 1 quarter of failure)				
0.06	0.03	0.90	0.69	0.58

Notes: This table reports the share of banks by capitalization, distinguishing by whether a bank is well, adequately, or under-capitalized; and if under-capitalized whether its significantly or critically under-capitalized.

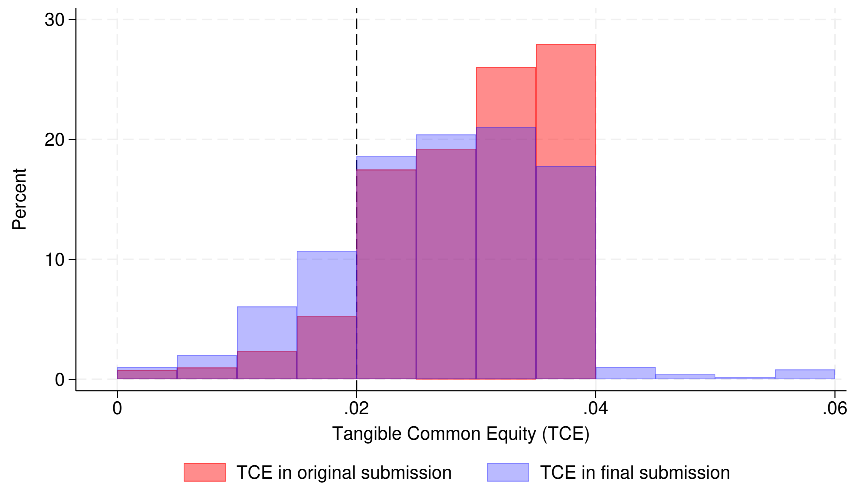
Call Report Revisions and Enforcement Actions Enforcement action is often formalized only when a bank is classified as either significantly or critically undercapitalized. This implies a potential interaction between supervisory audits—which require banks to restate their previously submitted financial statements—and legal supervisory intervention.

One way to understand the role of audits in helping pave the way for enforcement action is to study revisions around regulatory capital cutoffs that trigger drastic enforcement actions. For instance, a primary supervisor refers a bank to the FDIC, which in turn prepares a bank resolution when a bank becomes critically undercapitalized—that is, when its tangible common equity (TCE) falls short of 2% (see, e.g., Ivanov and Karolyi, 2023).[6]

To shed light on whether banks actively try to avoid crossing the cutoff that triggers this resolution process, Figure D.10 plots the distribution of TCE in banks with TCE less than 4% both before and after a resubmission of their Call Report during which income and equity to assets were restated. The figure suggests that there is a clear shift in banks' TCE moving below the 2 percent cutoff after resubmission. This shift of mass of banks with a TCE above 2% to banks with a TCE under 2% is in line with some banks taking advantage of their discretion in reporting financial statements, which allows them to maintain an official TCE ratio of above 2% and thus avoid triggering drastic action by the FDIC. However, following a supervisory on-site exam, some of these banks then report a TCE that crosses the threshold whereby the bank is referred to the FDIC. Thus, there is evidence of an interaction between the auditing function and enforcement function of supervisors, with on-site examinations leading to changes in banks' financial statements that pave the way for public enforcement action and resolution.

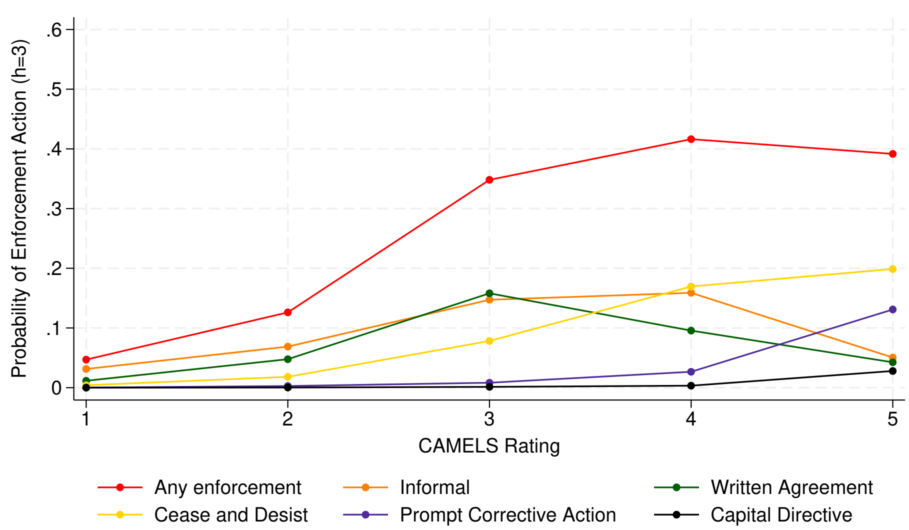
[6] See also the FDIC's resolution handbook.

Figure D.10: ***Call Report Revisions around the Prompt Corrective Action Cutoff***



Notes: This figure plots the distribution of Tangible Common Equity (TCE) as reported in both the first submitted and the final Call Report, conditional on a bank having TCE below 4% and conditional on the bank subsequently revising its original submission. The vertical line indicates the 2% cutoff at which a bank becomes formally "critically under-capitalized" and PCA is triggered.

Figure D.11: ***CAMELS Ratings and Enforcement Actions***



Notes: This figure plots the probability of enforcement actions from t to $t+4$ by CAMELS rating in t.