The Impact of Bank Examiners on Bank Risk Taking and Growth

Mehdi Beyhaghi, Jiyoung Chae, Filippo Curti, and Jeffrey R. Gerlach*

November 2025

Abstract

We study how individual bank examiners shape supervisory outcomes and bank behavior using supervisory ratings of U.S. bank holding companies from 2005–2023. We construct an examiner-specific effect (ESE)—the residual between actual and model-predicted ratings based on public data—to isolate heterogeneity in supervisory judgment. Examiners with greater access to private information deviate more from predicted ratings, while examiner stringency varies persistently and cyclically—tightening after crises and easing during expansions. Banks receiving tougher-than-expected ratings reduce asset growth and lending, whereas leniency has muted but risk-enhancing effects. These findings reveal that supervision is shaped not only by regulation but also by human discretion and information asymmetry, with implications for financial stability and the design of macroprudential oversight.

Keywords: Bank Supervision; Examiner Heterogeneity; Information Asymmetry; Regulatory Discretion; Bank Risk-Taking; Credit Allocation; Financial Stability.
JEL Classification: G21; G28; E58; D82.

*Beyhaghi is at the Board of Governers of the Federal Reserve System (e-mail: mehdi.beyhaghi@frb.gov), Chae is at the Federal Reserve Bank of Richmond (e-mail: jiyoung.chae@rich.frb.org). Curti is at the Federal Reserve Bank of Richmond (e-mail: filippo.curti@rich.frb.org). Gerlach is at the Federal Reserve Bank of Philadelphia (e-mail: jeffrey.gerlach@rich.frb.org). We thank Mark Flannery, Priyank Gandhi, John Kandrac, Fiona Hoellmann, Rahman Sakif, Lilah Shema-Zlatokrilov, seminar participants at the Federal Reserve Bank of Richmond, the Federal Reserve Board, as well as conference participants at the EFA annual meeting and the 2025 Behavioral Finance Conference at the University of Chicago for their comments and feedback. The views expressed in this article are solely those of the authors. They do not necessarily represent the views of the Federal Reserve Board, The Federal Reserve Bank of Richmond, the Federal Reserve Bank of Philadelphia, or the Federal Reserve System.

1 Introduction

In the United States, bank regulatory agencies allocate billions of dollars annually to supervision,[1] relying on thousands of bank examiners to ensure that banks adhere to regulations and operate safely and soundly.[2] This substantial investment reflects the critical importance of addressing a fundamental market friction in banking: information asymmetry.

In financial markets characterized by such asymmetric information, external monitoring mechanisms like bank examination serve a crucial certification role. Bank examiners function as privileged information producers who address opacity that markets alone cannot resolve. These examiners operate at the intersection of regulatory governance and financial intermediation, with their assessments directly influencing credit allocation, risk-taking, and ultimately, economic growth. Understanding how heterogeneity in examiner practices affects financial intermediation offers insights into fundamental mechanisms of information production in regulatory settings. This understanding is particularly important given the scale of resources devoted to bank supervision and banks' central role in supplying credit to individuals and businesses.

The outcomes of examinations, particularly the supervisory ratings, carry significant implications for banks. These can range from recommended corrective actions or mandated operational changes to revisions of financial statements. Some banks may face limitations on distributions or mergers and acquisitions, or they may even be forced to resolve. Examiners' assessments can thus have a significant impact on a bank's intermediation function and lead to tangible effects on the real economy. Although banking regulations are uniform, their enforcement can vary between examiners and even by the same examiner at different times, leading to differing assessments of the same bank.

In this paper, we investigate what explains heterogeneity in examiner practices and how this heterogeneity affects bank risk taking and lending. We find that examiners' assessments are influenced by key drivers beyond the bank's publicly observable fundamentals. These factors include examiners' access to private information, their individual attributes—such as relative stringency or leniency—and the prevailing macroeconomic conditions at the time

[1] According to a Congressional Research Service report from December 2020, the Federal Reserve reported spending $1.84 billion on supervision in 2019. Similarly, the Federal Deposit Insurance Corporation (FDIC) budgeted $1.06 billion for supervision in 2020, while the Office of the Comptroller of the Currency (OCC) spent $965 million on supervisory activities in fiscal year 2019. These are in addition to the state-level supervisory activities for state-chartered banks. Source: Congressional Research Service, "Banking Policy Issues in the 117th Congress," R46648, December 28, 2020, `https://crsreports.congress.gov.`

[2] According to estimates from Google AI, the three main federal agencies—the Federal Reserve, the Office of the Comptroller of the Currency (OCC), and the Federal Deposit Insurance Corporation (FDIC)—employ more than 7,000 examiners collectively.

of the examination. Furthermore, the heterogeneity in examiner practices has a significant impact on bank behavior. When a bank receives a tougher-than-predicted rating, it responds by reducing its risk-taking and scaling back investments. This adjustment holds even after controlling for bank and time fixed effects, as well as banks' underlying fundamentals.

Our methodology centers on developing a benchmark model that predicts supervisory ratings using data from 63,867 examinations conducted on 7,330 bank holding companies (hereafter referred to as "banks" for simplicity) from 2005 to 2023. While recent research, including the contemporaneous working paper by Agarwal, Morais, Seru, and Shue (2024), relies on bank-level ratings (CAMELS), our focus on bank holding companies provides a cleaner setting as the Federal Reserve is solely responsible for assigning the bank holding company rating, avoiding the confounding effects of multiple regulatory agencies and examination rotation present at the bank level. This model takes as input various risk measures and financial ratios derived from financial statements, as well as market data for publicly traded banks. The model provides an expected rating for each bank exam based on publicly available information. The difference between the actual rating and this expected rating represents the examiner-specific effect.

We then develop measures of access to private information and demonstrate that examiners with greater access to such information—that is, information beyond publicly available data—account for part of the difference between an examiner's actual rating and the model's predicted rating. Building on prior institutional-level research showing that proximity and on-site examinations facilitate information gathering by supervisory agencies (Gopalan, Kalda, and Manela, 2021; Kandrac and Schlusche, 2021), we examine these mechanisms at the individual examiner level. Specifically, we find that examiners who assess banks geographically closer to their office are more likely to deviate from the model-predicted rating. Conversely, examiners evaluating banks in regions different from their own are more likely to produce ratings that align closely with the model's predictions. Moreover, we find that examiners who conduct on-site examinations are more likely to produce ratings that differ from model-predicted ratings relative to those conducting off-site reviews. Examiners of publicly traded banks, which have more market information available, tend to produce ratings closer to model predictions, consistent with greater availability of public signals.

Our analysis also reveals that individual examiner characteristics, specifically their relative stringency or leniency compared to peers, play a significant role in rating outcomes. We measure examiner stringency as the average deviation of their ratings from expected ratings across their examinations of other banks, independent of the bank being assessed. Examiners identified as stringent based on past behavior are more likely to provide tougher-than-expected ratings, while lenient examiners tend to assign easier-than-expected ratings.

This pattern indicates that individual examiners bring their unique judgment and experience to each assessment. To address concerns that regulators may assign tougher examiners to weaker banks, we examine the relationship between a bank's previous rating and examiner stringency. We find no significant correlation, suggesting that assignments depend more on examiner availability than bank condition.

We then examine how examiner-specific effects evolve over time. We find that examiners' stringency varies systematically with the financial cycle—generally more lenient in periods preceding crises and more stringent in their aftermath. This pattern is evident during both the global financial crisis and the recent turmoil surrounding the failure of Silicon Valley Bank. These findings underscore that the timing of an exam can influence supervisory judgments, reflecting cyclical shifts in supervisory attitudes toward risk.

While our study was developed independently, a contemporaneous working paper by Agarwal, Morais, Seru, and Shue (2024) also examines examiner heterogeneity in supervisory ratings. Their analysis provides early evidence of personal discretion in bank-level examinations. Our paper, however, differs in focus, scope, and interpretation. We study bank holding companies—entities supervised exclusively by the Federal Reserve under a consistent framework—thus avoiding the institutional variation inherent in bank-level data. We also uncover systematic shifts in examiner behavior over time, showing that supervisory stringency tightens following crises and eases during expansions. Moreover, we provide evidence that part of the observed discrepancy between actual and model-predicted ratings reflects differences in examiners' access to private information rather than intrinsic examiner bias. Finally, we link examiner heterogeneity to real outcomes, documenting an asymmetric impact of examiner stringency on bank risk-taking and investment. Section 2 discusses in detail how our study complements and extends this work.

Our findings have important implications for understanding institutional factors affecting financial intermediation through the banking system. Regulatory agencies employ various strategies to minimize inconsistencies in examiner ratings, including standardized guidance, comprehensive training programs, supervisory review processes, and horizontal comparisons across similar banks. These efforts aim to ensure a consistent interpretation of supervisory expectations. Despite these measures, achieving perfect uniformity in a complex, evolving financial system remains challenging. This difficulty was evident in cases such as Silicon Valley Bank's failure in March 2023, where examiner ratings were notably more lenient than expected given the bank's financial condition and established supervisory norms (Barr, 2023).

More broadly, our results underscore the importance of supervision and the role of individual examiners as key actors in the information production process. Examiners differ in

experience, judgment, and information access, leading to variation in how supervision shapes bank behavior. Recognizing this heterogeneity is essential for understanding how regulation influences credit allocation and financial stability.

The remainder of the paper is structured as follows. Section 2 reviews related literature and our contribution. Section 3 outlines the institutional context. Section 4 develops our hypotheses. Section 5 describes the methodology. Section 6 presents empirical results. Section 7 concludes.

2 Related Literature

This paper contributes to research on external monitoring, particularly the role of examiners in bank supervision. A large literature highlights the inherent opacity of banks and the need for supervisory oversight. Dang, Gorton, Holmström, and Ordonez (2017) describe banks as "Secret Keepers," emphasizing their incentives to conceal loan-level information to prevent costly duplication of private information, while Morgan (2002) argues that the combination of unique assets, leverage, and systemic risk makes opacity an unavoidable feature of banking. Empirical evidence confirms this view: Berger and Davies (1998) show that unfavorable examination information predicts stock returns, and DeYoung, Flannery, Lang, and Sorescu (2001) find that supervisory ratings contain information not reflected in market prices, indicating that examinations complement market discipline.[3]

Bank supervision mitigates opacity and improves risk management by producing and acting on information unavailable to markets. Gopalan (2022) show that disclosure of supervisory ratings accelerates loan-loss provisioning; Costello, Granja, and Weber (2019) find that stricter regulators enhance reporting transparency; and Angelini, Lotti, Passalacqua, and Soggia (2021) and Bonfim, Cerqueiro, Degryse, and Ongena (2023) document that inspections reduce risk-taking and lending to weak firms. Correia, Luck, and Verner (2025) further show that supervisors use private information to anticipate failures. Collectively, these studies underscore that supervision complements market discipline by shaping bank behavior and limiting excessive risk.

Our paper builds on this foundation and contributes to the emerging literature on examiner heterogeneity in supervision. The closest to our work is the parallel study by Agarwal et al. (2024), who model confidential bank-level supervisory ratings (CAMELS composites) to quantify examiner discretion. While both studies examine examiner-specific effects, our analysis differs conceptually and empirically in several key dimensions.

[3]See also Hirtle (2016), Flannery, Kwan, and Nimalendran (2004), and Iannotta (2006) for further evidence on the informational value of supervision.

First, our approach to understanding rating dynamics is informed by key supervisory guidelines. The *Uniform Financial Institutions Rating System (UFIRS)* clarifies that "the composite rating generally bears a close relationship to the component ratings assigned, but it is not derived by computing an arithmetic average of component ratings."[4] This principle is reinforced in the FDIC's *Risk Management Manual*, which notes that "the composite rating is not derived by averaging component ratings; some components may be given more weight than others depending on the situation at an institution."[5] Similarly, the OCC's *Large Bank Supervision* handbook emphasizes that "the composite rating reflects the examiner's overall assessment of the bank's condition and risk profile, not a numeric average of the six factors."[6]

Guided by this regulatory framework, we recognize that differences between predicted and assigned composite ratings may reflect legitimate supervisory judgment based on qualitative factors—such as management responsiveness, governance quality, or emerging risks—not fully captured in quantitative data. Our methodology acknowledges that supervisory discretion represents intended flexibility within the rating system. We share the perspective advanced in the contemporaneous working paper by Agarwal et al. (2024) that persistent directional patterns across institutions may reveal an examiner's characteristic supervisory approach.

Our examination of examiner-specific effects considers that these patterns may reflect access to private or non-public information rather than solely representing discretionary judgment. Our specifications incorporating both examiner and bank fixed effects suggest that rating variation often corresponds with differences in information sets available during examination.

Second, our analysis centers on bank holding companies (BHCs) supervised exclusively by the Federal Reserve under the RFI/C(D) rating system.[7] This methodological choice allows us to examine entities evaluated under consistent supervisory standards. Additionally, because holding companies—not individual banks—are publicly traded entities, our approach enables us to connect supervisory assessments with market information, including equity valuations, debt spreads, and analyst forecasts. This integration of supervisory and market data enhances our ability to identify how examiner heterogeneity influences both regulatory outcomes and financial markets.

[4] https://www.federalreserve.gov/boarddocs/supmanual/cbem/200904/A50201.pdf

[5] https://www.fdic.gov/resources/supervision-and-examinations/examination-policies-manual/section1-1.pdf

[6] https://www.occ.gov/publications-and-resources/publications/comptrollers-handbook/files/large-bank-supervision/pub-ch-large-bank-supervision.pdf

[7] The literature documents cross-agency supervisory differences: Agarwal, Lucca, Seru, and Trebbi (2014) show that federal regulators are systematically tougher than state supervisors, and Beyhaghi and Gerlach (2024) extend these findings to a later period.

Finally, our paper introduces a time-series dimension. We show that examiner attitudes shift systematically over the business and financial cycle: heterogeneity intensifies before crises and moderates afterward, consistent with cyclical changes in supervisory risk tolerance. This dynamic perspective reveals that examiner stringency itself evolves with macro-financial conditions, offering new insight into how supervision adapts to risk environments.

A growing body of work emphasizes that differences in supervisors' access to information can fundamentally shape examination outcomes. Geographic proximity and examination modality directly affect the extent to which supervisors can observe and interpret bank conditions. Gopalan, Kalda, and Manela (2021) show that when supervisory offices close—thereby increasing distance between examiners and banks—institutions subsequently take on more leverage, distribute higher dividends, and face elevated failure risk, consistent with a decline in the quality of supervisory information. Similarly, Kandrac and Schlusche (2021) document that on-site examinations enable deeper information collection and materially influence supervisory ratings. These findings suggest that information asymmetries between banks and supervisors are not static but depend on institutional arrangements and the intensity of supervisory contact. Our paper builds on this perspective by identifying how such informational frictions manifest at the individual examiner level, showing that variation in proximity, examination type, and access to non-public information systematically influences examiner assessments within a unified supervisory framework.

Finally, our analysis connects to broader evidence on heterogeneity among external monitors such as auditors, analysts, and loan officers. Francis and Yu (2009) and Gul, Jaggi, and Krishnan (2007) document that audit quality varies with office size, expertise, and incentives, while Clement (1999) and Ke and Yu (2006) show that analysts' forecast accuracy depends on individual ability and information access. In banking, Berger, Kick, and Schaeck (2014), Srivastav and Hagendorff (2016), and Cole, Kanz, and Klapper (2015) find that governance structures and loan-officer incentives shape risk-taking. These studies, like ours, highlight that the identity of the monitor matters even within formalized oversight frameworks.

3 Institutional Background

From 2005 to 2019, the Federal Reserve has used the RFI/C(D) rating system (RFI for short) to communicate its supervisory assessment of all bank holding companies regardless of their asset size, complexity, or systemic importance.[8]

[8]R, F, I, C, and D in RFI/C(D) respectively represent risk management, financial condition, impact, composite, and [subsidiary] depository institutions. Bank holding companies receive a rating ranging from 1 to 5 for the composite rating and each of its sub-component, with lower numbers denoting fewer issues and thus a better rating. For more details see Federal Reserve (2019).

The RFI rating system provides an assessment of risk management and financial condition factors that are common to holding companies. Under this system, the Federal Reserve tries to ensure that applicable bank holding companies are evaluated in a comprehensive and uniform manner, and that supervisory attention is appropriately focused on the holding companies that exhibit financial and operational weaknesses or adverse trends. The RFI rating system is a crucial tool for regulators. It helps identify problems with bank holding companies early on. It also enables the Federal Reserve to track safety and soundness trends and assess the overall strength of the financial industry.

Under the RFI rating system, each bank holding company is assigned a composite rating (C) based on an overall evaluation of its managerial and financial condition. This composite rating, which we use as our main dependent variable in our analyses, is derived from several components. These include evaluations of the bank's risk management practices, financial condition, and the interplay between its depository and non-depository subsidiaries. According to examiner manuals, the composite rating encompasses both a forward-looking and static assessment of the consolidated organization, as well as an assessment of the relationship between the depository and nondepository entities. The composite rating is not derived as a simple numeric average of each components; rather, it reflects examiner judgment with respect to the relative importance of each component to the safe-and-sound operation of the holding company. The composite rating is assigned on a numeric scale from 1 to 5, with 1 indicating the highest and 5 representing the lowest rating. Ratings of 1 or 2 signify a satisfactory state with minimal supervisory issues, whereas ratings of 3, 4, or 5 indicate escalating levels of regulatory concerns, ranging from moderate to extreme (Federal Reserve (2019)).

Moreover, Federal Reserve examiners assess bank holding companies through a process that can involve on-site visits and/or off-site monitoring. The Federal Reserve System comprises 12 districts, each overseen by a Federal Reserve Bank responsible for supervising bank holding companies in its region. These banks operate from headquarters and up to five branches located in various cities within their district.[9] Examiners in each branch typically examine the banks that are geographically closer. A key component of this supervision is the annual safety and soundness examination, which typically results in the assignment of an RFI rating.

[9]The list of Federal Reserve Banks and their branches includes: Boston (no branches); New York (no branches); Philadelphia (no branches); Cleveland (Cincinnati and Pittsburgh); Richmond (Baltimore and Charlotte); Atlanta (Birmingham, Jacksonville, Miami, Nashville, and New Orleans); Chicago (Detroit); St. Louis (Little Rock, Louisville, and Memphis); Minneapolis (Helena); Kansas City (Denver, Oklahoma City, and Omaha); Dallas (El Paso, Houston, and San Antonio); and San Francisco (Los Angeles, Portland, Salt Lake City, and Seattle).

Further, while the RFI system remains the dominant rating system for bank holding companies in the US after 2019, a new rating system was introduced in November 2018 and became effective in February 2019. This new system, known as the LFI (Large Financial Institution) system, applies to bank holding companies with $100 billion or more in assets. Fewer than 1% of banks are subject to this new system.[10] Prior to RFI ratings (i.e., before 2005), the Federal Reserve used the BOPEC rating system for all bank holding companies. BOPEC stands for five key areas of supervisory concern: the condition of the bank holding company's bank and nonbank subsidiaries, Parent company, Earnings, and Capital adequacy. Bank holding companies with the best performance are assigned a BOPEC rating of one, while those with the worst performance are given a BOPEC rating of five. The RFI rating system deviates from BOPEC rating system in that unlike BOPEC, which mostly relies on historical analysis of bank holding company financial conditions, RFI relies on a forward-looking assessment of risk management and financial factors (Krainer and Lopez, 2009).

4 Hypothesis Development

As explained while banking regulations are applied uniformly across banks, regulatory agencies rely on individual bank examiners to enforce these regulations and supervise banks. Banks cannot influence which examiner will be assigned to them; examiners are typically assigned based on availability. Each examiner generally examines multiple banks during their tenure, sometimes examining multiple banks in a single year. Conversely, a typical bank may be examined by multiple examiners over time.

As explained before, bank examiners are responsible for conducting periodic examinations to assess banks' financial conditions and regulatory compliance. They identify and evaluate various risks associated with bank operations, recommend corrective actions, and provide guidance on best practices. Examiners prepare reports on their findings and recommendations, which are shared with Federal Reserve officials as well as banks' senior management and boards of directors. Examiners' assessments are crucial for ensuring the safety and soundness of the financial system.

As mentioned in the introduction, bank regulatory agencies take various measures to standardize examination practices. These include providing standardized guidance to examiners, implementing training programs, and conducting supervisory review processes and horizontal comparisons of ratings assigned to banks. However, due to the complexity of banks

[10] At the end of 2023, there were 3,407 domestic bank holding companies, of which 23 held more than $100 billion in assets. See Bank Holding Companies: Background and Issues for Congress (December 5, 2024) available at `https://crsreports.congress.gov/`.

and the unique nature of each bank's operations and regional context, perfect consistency is not achievable.

Examiners have access to banks' financial statements and other publicly available data about the bank, but their ability to access and process private information can vary due to factors beyond their control. Onsite examinations allow for direct interviews with bank staff and managers, facilitating the collection of private information, while offsite examinations may rely more heavily on bank-provided documentation. The geographical arrangement of examinations also impacts information access. For instance, the San Francisco Fed oversees a vast district spanning nine western states, including Alaska and Hawaii, with branches in Los Angeles, Portland, Salt Lake City, and Seattle. An examiner based in San Francisco conducting an onsite examination of a bank in Anchorage, Alaska, would need to travel over 2,000 miles, potentially limiting their access to local information. Conversely, an examiner living in San Francisco may have more insight into banks located in San Francisco. They are likely more familiar with local news and economic conditions, and may even bank with or have worked closely with local banks, providing them with deeper knowledge about particular banks in their region. These factors contribute to variations in examiners' access to and processing of information.

Hypothesis 1: Examiners with greater access to private information are more likely to assign ratings that differ from those based solely on public information.

Further, examiners, as individuals, bring their own unique perspectives and traits to the examination process. These personal characteristics can potentially influence their assessments and ratings of banks. Some examiners may be inherently more conservative in their risk assessments, leading them to be consistently tougher in their evaluations. These examiners might place greater emphasis on potential risks and vulnerabilities, resulting in more stringent ratings and recommendations. Conversely, other examiners may adopt a more lenient approach, perhaps giving greater weight to mitigating factors or potential upsides in a bank's operations. This variation in examiner toughness can persist even when presented with identical information about a bank. Factors contributing to these differences may include an examiner's professional background, personal risk tolerance, and their individual interpretation of regulatory guidelines. As a result, two examiners evaluating the same bank with access to the same information may arrive at different conclusions and assign different ratings, reflecting the inherent subjectivity in the examination process despite efforts to standardize procedures.

Hypothesis 2: Individual examiner characteristics, particularly relative strictness or leniency, influence bank examination outcomes, leading to differences in ratings across examiners, even when controlling for bank-specific factors and economic conditions.

Economic conditions and recent financial events can also significantly influence examiner judgments. Following a banking crisis or financial turmoil, examiners often adopt a more stringent stance in response to increased public scrutiny and regulatory pressure. This can lead to tougher examinations and more conservative risk assessments, even for banks not directly involved in the crisis. Conversely, during periods of economic stability and growth, examiners might become more lenient, potentially giving more weight to recent positive performance. This pattern suggests a relationship between economic conditions and the rigor of bank examinations, with stricter oversight often coinciding with economic downturns and more relaxed approaches during expansionary periods.

Hypothesis 3: The average stringency of bank examinations can vary over time, with examiners tending to be more stringent following periods of financial stress or economic downturns, and relatively more lenient during periods of economic stability and growth.

Lastly, bank examination ratings carry significant implications, particularly when they indicate examiners' dissatisfaction with a bank's situation. These ratings often include specific corrective actions and are linked to potential consequences, such as enforcement actions or restrictions on high-risk activities, which can significantly shape a bank's strategic decisions, operational practices, and overall risk profile. Ratings also serve as a crucial communication tool, conveying supervisory concerns to bank management and boards of directors. This communication influences resource allocation within banks and plays a direct role in risk mitigation.

A stringent rating typically compels banks to reduce their risk exposure, enhancing the institution's safety and stability. However, this risk reduction often comes at the cost of limiting growth, cutting asset expansion, and curtailing lending activities. For instance, a bank receiving a poor rating might need to increase capital reserves, reducing available funds for loans, or divest from high-risk, high-return assets. While these actions can strengthen the bank's financial position, they may simultaneously slow its growth trajectory and limit market share expansion. Thus, examination ratings highlight the delicate balance between ensuring financial stability and fostering economic growth through bank lending and invest-

ment activities.

Hypothesis 4: Banks receiving more stringent examination ratings will exhibit reduced risk-taking behavior but also experience slower growth compared to banks with less stringent ratings, all else being equal.

5 Methodology

To test our hypotheses, we employ a two-stage residual inclusion model, based on the work of Hausman (1978). We conceptualize a bank's rating as comprising two main components. The first is an objective assessment of bank risk, derived from publicly available information about the bank, including financial statements. The second is an examiner-specific effect, which encompasses both the examiner's access to private information about the bank and the examiner's inherent toughness or leniency in assigning ratings. The private information aspect is related to the bank's actual condition but not captured in public data, the examiner's inherent toughness or leniency is orthogonal to the bank's actual condition.

To isolate the examiner-specific effect, we conduct a regression analysis using data from all available examinations. In this model, we regress bank ratings on various measures of bank risk derived from financial statements and market data (for publicly traded banks). These measures encompass a wide range of financial indicators, including various aspects of capital, liquidity, performance, volatility, concentration, business model, and asset and liability compositions. The residual from this regression—the unexplained variation in ratings—is assumed to represent the examiner-specific effect. We then investigate this residual to explore its relationship with our constructed proxies for examiner access to private information and examiner stringency. Furthermore, we utilize this examiner-specific effect as an independent variable in second-stage regressions, to assess its impact on subsequent bank growth or risk-taking behavior, while controlling for objectively measurable bank characteristics.

More specifically, we implement the following steps:

- Step 1: Estimating the part of the bank rating that is not related to the model-predicted rating based on publicly available information.

  Using Ordinary Least Squares (OLS) specification:

$$R_{i,t} = \alpha + \beta \text{Bank Characteristics}_{i,t} + \epsilon_{i,t} \tag{1}$$

  Using Ordered Logit specification:

$$\text{logit}(P(R_{i,t} \leq j)) = \kappa_j - \beta \text{Bank Characteristics}_{i,t}, \quad j = 1, 2, 3, 4 \tag{2}$$

Using Logit specification:

$$P(BR_{i,t} = 1) = \frac{1}{1 + e^{-(\alpha + \beta \text{Bank Characteristics}_{i,t})}} \tag{3}$$

where $R_{i,t}$ represents the rating produced by an examination of bank i at time t (quarters). $R_{i,t}$ is an ordinal rating variable (e.g., 1, 2, 3, 4, 5). For the Logit specification the dependent variable $BR_{i,t}$ is a binary variable that takes the value of 1 if rating is 3, 4, or 5 and zero if the rating is 1 or 2. β is the vector of coefficients. Further, in the OLS regression, α is the intercept and $\epsilon_{i,t}$ represents the rating shock. In the Ordered Logit regression j represents the possible rating values (1 to 4 in the first equation, as the probability of being less than or equal to 5 is always 1) and κ_j are the cutpoint parameters.

As will be discussed we compare estimation under these specifications, but use OLS as our main specification going forward.

- Step 2: Constructing "examiner-specific effect" as

$$ESE_{i,t} = R_{i,t} - Round(\widehat{R}_{i,t}) \tag{4}$$

which is the difference between the actual rating and the model-predicted rating, rounded to the nearest integer. Bank ratings are assigned as integers ranging from 1 to 5. To be consistent we round the predicted rating to the closest integer. This approach also helps to isolate more substantial deviations that are more likely to represent genuine examiner-specific effects.

- Step 3: Investigating the relationship between the examiner-specific effect and proxies of access to private information.

$$|\text{ESE}_{i,t}| = \beta_1 \text{PAIP}_{i,t} + \beta_2 (\text{Controls})_{i,t} + \delta_i + \mu_t + \eta_{i,t} \tag{5}$$

where $\text{PAIP}_{i,t}$ is the proxy for access to private information in the examination of bank i at time t, and δ_i and μ_t, respectively, represent bank and time fixed effects. We cannot directly observe the nature or direction of private information available to examiners. However, as specified in Hypothesis 1, we expect that ratings assigned by

examiners with greater access to private information are more likely to deviate from the model-predicted ratings. To capture this effect, we use the absolute value of the examiner-specific rating as the dependent variable in our analysis.

Our proxies for access to private information include various measures of distance between the examiner's office and the bank's location, whether the examination was conducted onsite or offsite, and whether the bank is publicly traded. We assume that examiners gain more private information when they are geographically closer to the bank or when conducting onsite examinations. Further, we assume that examiners have more private information relative to markets when the bank is not publicly traded.

- Step 4: Assessing whether the examiner specific effect is related to a proxy of examiner stringency.

$$\text{ESE}_{i,t} = \gamma_1 E_{k,i} + \gamma_2(\text{Controls})_{i,t} + \delta_i + \mu_t + \nu_{i,k,t} \tag{6}$$

where k denotes the examiner that examines bank i at time t. $E_{i,k}$ is the proxy for examiner stringency, calculated as

$$E_{k,i} = \frac{1}{Nk,-i} \sum_{j \neq i} \text{ESE}_{j,t(k,j)} \tag{7}$$

where $E_{k,i}$ is the stringency of examiner k for examining bank i. $\text{ESE}_{j,t(k,j)}$ is the examiner specific effect for bank j at the time t when examiner k examined bank j. $Nk,-i$ is the number of examinations conducted by examiner k for all banks except bank i. This formula calculates the average Examiner Specific Effect across all examinations conducted by examiner excluding the examination of the current bank.[11]

- Step 5: Assess the impact of macroeconomic conditions on examination stringency. We calculate annual averages of examiner-specific effects across all examinations for each year. We plot these averages over time with confidence intervals. If the intervals exclude zero, it indicates statistically significant shifts in overall examination stringency. We attribute temporal fluctuations in average examiner effect primarily to changes in examiner stringency rather than sudden variations in access to private information. This approach allows us to identify temporal patterns in stringency and their potential correlation with macroeconomic events, such as financial crises.

[11]This measure of examiner stringency is consistent with a leave-one-out approach, similar to jackknife estimation methods in statistics. By excluding the current bank from the calculation, we reduce potential bias and provide a more robust estimate of the examiner's overall stringency.

- Step 6: Assessing whether the residual obtained from Step 1 is predictive of the bank's future risk taking and asset growth.

$$Y_{i,t+1} = \lambda_1 \text{ESE}_{i,t} + \delta_i + \mu_t + \xi_{i,t} \tag{8}$$

$$Y_{i,t+1} = \lambda_2 \text{ESE}_{i,t} + \delta_i + \mu_t + \rho_k + \zeta_{i,k,t} \tag{9}$$

where $Y_{i,t+1}$ represents a measure of risk-taking such as capital ratio or growth such as total assets or ratio of loans to assets for bank i in year $t+1$.

In the second regression, we incorporate examiner fixed effects to control for individual examiners' inherent stringency levels. By doing so, we aim to isolate the impact of examiners' access to private information on their assessments. Our hypothesis is that after controlling for examiner-specific stringency, any remaining variation in ratings can be attributed to differences in access to private information.

6 Empirical analysis

6.1 Data

Our empirical analysis draws from three main data sources, with the Federal Reserve's confidential supervision records serving as the primary source. These records provide examination information and supervisory ratings, specifically the RFI composite ratings of bank holding companies (banks). Our dataset encompasses 25,245 annual examinations of 2,593 unique banks over the period 2005–2023, with 2005 marking the first year of RFI rating enforcement. We include all available ratings for firms that filed FR Y-9C at least once during the sample period, even for periods when their total assets were below the $3 billion FR Y-9C filing threshold. The RFI composite rating of 1 indicates strong performance, 2 satisfactory, 3 fair, 4 marginal, and 5 unsatisfactory. Table 1 presents the distribution of these ratings across our sample. As can be seen, the majority of banks are rated satisfactory.

Our second data source comprises financial reports from bank holding companies, primarily FR Y-9C filings. Many banks, however, alternate between FR Y-9C and FR Y-9SP filings due to fluctuations around the $3 billion total asset threshold. As explained, to maintain comprehensive coverage, we incorporate measures from both report types. While FR Y-9SP filings are simpler and less frequent, lacking detailed metrics like Tier 1 capital, our examination rating data remains consistent across all filers. To address the varying depth of financial information, we construct two sets of explanatory variables: a baseline set using

measures common to both statement types, and an expanded set derived solely from FR Y-9C data. This dual approach maximizes the use of information while utilizing the more detailed information available for larger institutions when possible.

The summary statistics of the variables used in our analysis are provided in Table 2. The first six variables (total assets and its logarithm, as well as ratios of equity to assets, securities to assets, loans to assets, and ROA) constitute our baseline set, which can be constructed for all observations regardless of filing type. These variables were selected based on prior literature and broadly capture bank size, capital, asset distribution, and profitability—factors expected to forecast ratings effectively. The subsequent 20 variables are derived exclusively from FR Y-9C filings, offering a more detailed financial profile of the larger institutions.[12] These additional variables are used as regressors in our expanded forecasting models. The financial data are taken from the most recent statements available prior to the bank examination.

Our third source of data is market data from the Center for Research in Security Prices (CRSP) for publicly traded banks. We obtain data on equity returns, prices, and shares outstanding using the CRSP-FRB Link dataset, which is provided by the Federal Reserve Bank of New York.[13] This dataset establishes a connection between CRSP's permanent company number (PERMCO) and the unique RSSD ID number of the bank holding company. We utilize market information up to the end of the month preceding the end of the examination period, which is before the assignment of a rating. For instance, if a rating is assigned on May 10, we consider all stock market information up until April 30. Consequently, we derive the following metrics: market-to-book ratio and average monthly stock return over the most recent quarter (rolling 12-month period), as well as average monthly stock return, standard deviation of monthly returns, and average turnover (volume/shares outstanding) over a rolling 12-month period. The summary statistics for these variables are presented at the end of Table 2.

These variables are defined in Table A1 in the Appendix. All ratios are winsorized at the 3rd and 97th percentiles to mitigate the effect of outliers.

[12]These include the following ratios: trading assets/assets, allowance for loans/assets, leverage ratio, double leverage ratio, tier 1 capital ratio, ratio of residential real estate loans, ratio of commercial real estate loans, ratio of commercial and industrial loans, ratio of consumer loans, ratio of depository institution loans, ratio of foreign institution loans, ratio of agriculture loans, loans Herfindahl-Hirschman Index, assets Herfindahl-Hirschman, risk-weighted assets/assets, NPL/assets, NPL/loans, loan loss reserve/loans, loan loss reserve/NPL, residential real estate NPL, commercial real estate NPL, commercial and industrial NPL, consumer NPL. Definitions are provided in Table A1 in the Appendix.

[13]`https://www.newyorkfed.org/research/banking_research/datasets.html`

6.2 Estimating Examiner-Specific Effect

We initiate our analysis with a forecasting model that predicts bank ratings using publicly available information, as outlined in Step 1 of our methodology. Starting with a parsimonious set of bank characteristics derived from FR Y-9C or FR Y-9SP statements, we progressively expand our model before proceeding with subsequent steps. Our findings are presented sequentially, aligned with our hypotheses.

Our baseline forecasting model is an Ordinary Least Squares (OLS) regression using six bank-year level characteristics to predict ratings across all data. Notably we use last/starting composite rating as the first variable, which is also an input to the examiner's rating. The other five are the variables we extract from financial statements. Given the discrete nature of bank ratings (1-5), we also employ alternative discrete choice models as per step 1 of our methodology. These include an ordered logit model and a logistic model. The dependent variable for the latter is binary (1 for ratings 3-5, 0 for ratings 1-2). We compare these models to assess their relative performance in forecasting ratings. Results are presented in Table 3.

Table 3 shows consistent coefficient signs and significance across all three specifications. The OLS model's adjusted R-squared of 0.725 indicates that it explains 72.5% of the variation in bank ratings. We use the residuals from this model as a measure of examiner-specific effects in subsequent analyses.

It's important to note that this is a forecasting model, and the coefficients should be interpreted with caution due to potential correlations among variables. In forecasting models, we aim to incorporate as many relevant variables as possible to provide the most accurate prediction. The current model uses bank characteristics derived from both FR Y-9C and FR Y-9SP filings, which we later augment with additional variables in an expanded version when we focus on FR Y-9C banks.

In Figure 1 we compare the performances of the OLS and the Ordered Logit model fits in predicting bank ratings. The figure illustrates the distribution of examiner-specific effects computed using both models. In Figure 1, we observe that both models show a similar concentration of examiner-specific effects around zero. This suggests that both models have a similar level of accuracy in predicting the actual ratings assigned by examiners. Given this, we use the examiner-specific effects derived from the OLS model in all our subsequent analyses.

To further explore the model's predictive power, we focus on the subsample of publicly traded banks in Table 4. This subset allows us to access a broader range of regressors. Column 1 presents our baseline model with six explanatory variables, achieving an adjusted

R-squared of 0.676. In Column 2, we expand to 29 explanatory variables include starting composite rating and the 28 FR Y-9C specific variables that were described in the data section. Including these regressors increases the adjusted R-squared to 0.702. Column 3 further augments the model by adding market value metrics that represent stock return, volatility and volume, bringing the total to 34 variables and marginally improving the adjusted R-squared to 0.710. These variables were also described in the data section. The results from Table 4 demonstrate that additional variables offer incremental improvements to the model's forecast. They also show that our parsimonious baseline model performs relatively well in predicting bank ratings. The modest gains in predictive power from the expanded models underscore the effectiveness of our initial variable selection.

Following this step, we assign measures of the examiner-specific effect to each bank-year observation based on the maximum amount of available information available for each observation. That means, for publicly traded banks, we utilize the saturated model incorporating both financial statement data and additional market variables. FR Y-9C filers that are not publicly traded are assessed using the saturated model with FR Y-9C variables, while the remaining banks are evaluated using our baseline model. This stratified approach maximizes the use of available information for each bank type.

6.3 Testing H1 - Access to Private Information and Rating Divergence

To test Hypothesis 1, we analyze how examiners' access to private information affects the deviation of their ratings from model predictions based on public data. Results for this analysis are presented in Tables 5 and 6.

As outlined in the Methodology section, we employ three proxies for access to private information: examiner-bank distance, examination type (off-site vs. on-site), and bank public trading status. Table 5 presents results for distance and examination type, while Table 6 focuses on publicly traded banks. The dependent variable in these Tables is the absolute examiner-specific deviation.

In Table 5, we consider two different ways to capture access to private information based on distance. The first is a simple measure with a dummy that takes a value of one if the distance between bank and examiner's office is more than the median in the sample (Models 1, 3, and 7). The second uses three dummies based on whether the distance between bank and examiner office is less than 25 miles, between 25 miles and 50 miles, and more than 50 miles, following DeYoung, Glennon, and Nigro (2008) (Models 2, 4, and 8, with the omitted variable being the less than 25 miles dummy).

Consistent with Hypothesis 1, we find that examiners' ratings for banks further from their offices are more likely to align with model-predicted ratings based on public information. Specifically, being above median distance, reduces examiner's rating deviation by 2.1 percentage points (Model 1), or by 10% of the average deviation of 0.219. Further, being 50+ miles away reduces the deviation by 3.3 percentage points (Model 3) or by 15% of average deviation in the sample. The results are robust to controlling for time fixed effects. Overall the results show that examiners that are examining a bank further away from their office are more likely to produce ratings that are closer to predicted ratings based on publicly available information.

Furthermore, our other measure of access to private information is a dummy variable that takes a value of one if the exam is conducted off-site. Importantly, while there is a positive correlation between a bank's distance from the examiner's office and the likelihood of an off-site exam, we find that the off-site dummy provides additional explanatory power even after controlling for distance variables. In Model 7, the coefficient on the off-site dummy is -0.027, significant at the 1% level, suggesting that off-site examinations are 12% more likely (relative to mean deviation) to produce ratings aligned with model predictions. Model 8 demonstrates that when a bank is at least 50 miles away from the examiner's office and the exam is conducted off-site, the examiner's rating is 7 percentage points closer to the model-predicted rating compared to the alternative. This reduction is equivalent to one-third of the average deviation.

Table 6 provides additional support for Hypothesis 1 by demonstrating that examiner estimates for publicly traded banks exhibit smaller deviations from model predictions compared to other banks. The table presents four different models, each varying in their choice of year fixed effects and controls. Notably, all models consistently show a significant negative relationship between the publicly traded dummy variable and the absolute examiner-specific deviation.

In the most restrictive specification (Model 4), the coefficient on the publicly traded dummy is -0.042, statistically significant at the 1% level. This indicates that public trading status reduces the discrepancy between examiner ratings and model predictions by 4.2 percentage points. Considering the average deviation of 0.219, this reduction is equivalent to approximately 20% of the average deviation.

This finding aligns with our hypothesis that publicly traded banks are subject to greater scrutiny from various market participants, including stock analysts, investors, and regulatory bodies such as the SEC. This increased information production and transparency effectively diminishes the examiner's private information advantage, resulting in ratings that more closely align with model predictions based on publicly available information.

Overall, the results in Tables 5 and 6 are consistent with Hypothesis 1, demonstrating that parts of heterogeneity in examiners' ratings can be attributed to the examiners' access to private information.

6.4 Testing H2 - Examiner Stringency and Rating Consistency

Next, we examine the role of individual examiner characteristics, particularly, examiner's relative stringency on rating assignments. Our analysis of Hypothesis 2, presented in Table 7, reveals that individual examiner characteristics significantly influence bank ratings, even after controlling for bank-specific factors and economic conditions. In the first five models of Table 7, we use a variety of controls to check the robustness of our findings. Our key explanatory variable is the measure of examiner stringency that is calculated as the average discrepancy between the examiner's ratings and model-predicted ratings for other banks the examiner examine. The coefficients for this variable are significant at the 1% level across all specifications, including those with time and bank fixed effects.

In the most restrictive model (Model 5), the coefficient for examiner stringency is 0.246. This indicates that an examiner who has, on average, given ratings one grade worse than the model-predicted rating in examinations of other banks is 24.6% more likely than other examiners to give a rating one grade worse than the model-predicted rating for the current bank examination. In other words, examiners who have been tough in previous examinations are more likely to be tough in current examinations as well. These results are robust when accounting for firm- and time-fixed effects, as well as the choice of bank controls.

A potential concern regarding our results is the possibility of selection bias in examiner assignment, specifically, if tougher examiners are systematically assigned to weaker banks—those that are in a more precarious financial position than their public data or financial statements would indicate. If this were the case, it could lead to a spurious correlation between examiner stringency and bank outcomes, potentially biasing our results. To address this concern, we implement an additional test in the final specification of Table 7. This test involves regressing the starting rating of the bank on examiner stringency. The rationale behind this approach is that if tougher examiners were indeed systematically assigned to weaker banks, we would expect to see a significant positive relationship between examiner stringency and the initial bank rating.

Our results, however, do not support this alternative explanation. The coefficient on examiner stringency in this regression is not statistically different from zero, which suggests that there is no systematic relationship between examiner stringency and the initial condition of the banks they are assigned to examine. This finding aligns with our understanding of

the examiner assignment process, which is primarily based on examiner availability rather than an intentional matching of examiner characteristics to bank conditions.

Overall, we the results in Table 7 support Hypothesis 2 in that the variations we observe in bank outcomes following examinations are partially attributable to differences in examiner stringency. In other words, examiners' intrinsic characteristics can influence their ratings.

6.5 Testing H3 - Temporal Trends in Examiner Stringency

Next, we move to test Hypothesis 3, which posits that the average stringency of bank examinations varies over time, with examiners tending to be more stringent following periods of financial stress or economic downturns, and relatively more lenient during periods of economic stability and growth. To investigate this hypothesis, we plot the annual average examiner-specific effects in Figure 2.

Figure 2 shows the temporal variation in examiner stringency. The figure plots the mean examiner-specific effects with 90% confidence intervals for each year, where a negative value indicates that ratings in that year were on average more lenient than expected, and a positive value indicates that ratings are on average more stringent than expected. The shaded areas on the figure represent the Global Financial Crisis (2007Q4-2009Q2) and the banking turmoil of 2023Q1.

Our results reveal clear patterns of leniency during periods leading up to financial crises and increased stringency following crises. For instance, during the Global Financial Crisis, the average examiner-specific effect reached its lowest level (-0.280), indicating a period of relative leniency, and subsequently rebounded to its highest level (0.080) in 2014, following the crisis, as examiners responded to increased public scrutiny and attention from regulators, as well as the enactment of the Dodd-Frank Act, by adopting a more stringent approach to bank examinations. We also observe that bank ratings tend to be relatively lenient in the years leading up to the banking turmoil of 2023 that led to the failure of Silicon Valley Bank, Signature Bank and First Republic Bank following a period of economic expansion. Overall, the results in this section support our hypothesis that the average examiner is possibly influenced by the macroeconomic environment.

The results so far support our Hypotheses 1-3, demonstrating that factors such as access to private information, examiner intrinsic stringency, and changing economic conditions contribute to heterogeneity in examiner practices and affect bank ratings. In the next set of tests, we examine how these examiner-specific effects are related to future bank risk-taking and growth, addressing Hypothesis 4.

6.6 Testing H4 - Impact of Examiner-Specific Effects on Bank Behavior

To analyze Hypothesis 4, we investigate the relationship between examiner-specific effects and future bank behavior. We use our main measure of examiner-specific effect (ESE_t),[14] as well as two decomposed measures to explore potential asymmetric effects: a positive component (ESE_t^+), capturing tougher-than-predicted ratings, and a negative component (ESE_t^-), capturing more lenient ratings. Formally, $ESE_t^+ = \max(ESE_t, 0)$ and $ESE_t^- = \max(-ESE_t, 0)$. To ensure interpretive consistency, we exclude from the construction of ESE_t^+ those bank-year observations where the actual rating is 1, as such observations cannot reflect stricter-than-predicted outcomes.[15] Similarly, we exclude actual ratings of 5 in ESE_t^+ regressions.

We estimate panel regressions with bank and year fixed effects to evaluate how these measures relate to bank behavior in the subsequent year. Table 8 presents results using our baseline specification for Log(Assets) in Model 1. Due to evidence of asymmetric responses to examiner-specific effects across all dependent variables, we primarily report results for ESE_t^+ and ESE_t^- in Models 2–11 to conserve space. Our findings generally indicate that banks respond more strongly to tougher-than-predicted ratings (ESE_t^+). This is evidenced by statistically and economically significant effects on key metrics such as bank size, lending activity, and capital ratios. These results suggest that examiner stringency has a more pronounced impact on future bank behavior compared to examiner leniency.

Notably, in Model 2, a one-grade increase above the model-predicted rating is associated with approximately a 5.9% reduction in total assets and a 0.6 percentage point decrease in the loan-to-asset ratio (compared to a sample mean of 34.5%). Moreover, we observe a positive relationship with bank capital. Specifically, a one-grade higher than predicted rating correlates with a 0.3 percentage point increase in the equity-to-assets ratio and a 0.4 percentage point increase in the Tier 1 ratio. These effects are substantial when compared to the sample averages of 9.7% for equity-to-assets and 13.1% for the Tier 1 ratio. These findings strongly support Hypothesis 4, demonstrating that banks receiving tougher-than-predicted ratings subsequently reduce both risk-taking and growth. The observed changes in asset size, lending activity, and capital ratios collectively indicate a more conservative approach to banking operations following stricter supervisory assessments.

[14] ESE_t represents $ESE_{i,t}$ for brevity. All regressions are conducted at the bank(i)-year(t) level, consistent with our previous analyses.

[15] A rating of 1 is already the most favorable supervisory rating and cannot be stricter than expected. Similarly, for the construction of ESE_t^-, we exclude observations with an actual rating of 5, which is the harshest possible rating. These constraints ensure that the decomposed measures capture meaningful upward and downward rating deviations.

In contrast to the pronounced effects of stricter ratings, our analysis reveals that responses to lenient ratings (ESE_t^-) are generally muted or statistically insignificant across most metrics. However, there is one notable exception: the non-performing loans (NPL) ratio. We find that lenient ratings are associated with a statistically significant increase in NPLs, suggesting that supervisory leniency may allow risk to accumulate within banks. Quantitatively, a rating that is one grade more lenient than the model-predicted rating corresponds to a 0.1 percentage point increase in the NPL-to-asset ratio. This effect is meaningful when compared to the sample average NPL-to-assets ratio of 1.3%. These findings provide nuanced support for the second part of Hypothesis 4, indicating that while banks may not significantly alter their overall behavior in response to lenient ratings, there is evidence of increased risk accumulation in their loan portfolios.

Table 9 replicates the analysis from Table 8 while incorporating examiner fixed effects. This approach aims to isolate the component of the examiner-specific effect related to private information, net of persistent examiner-level toughness or leniency. Notably, our results remain robust under this specification. For instance, the effect on asset growth is -0.056, nearly identical to the baseline model. This consistency suggests that a significant portion of the effect is driven by private, non-public information uncovered during the examination process.

Our findings continue to support an asymmetric response pattern. Tougher-than-expected ratings consistently drive contraction and de-risking across various metrics. In contrast, lenient ratings show less systematic effects, primarily manifesting in higher NPLs and, in some instances, a modest increase in total assets.

These results not only corroborate Hypothesis 4 but also carry substantial regulatory implications. While uniform regulations guide bank supervision, our analysis reveals that examiner-level heterogeneity—particularly access to private information—can significantly influence bank behavior.

Overall, our evidence demonstrates that examiner-specific effects are not mere statistical artifacts but meaningful supervisory signals. Ratings that deviate from model predictions based on public data have tangible consequences for banks' balance sheet strategies and risk postures in the subsequent year. These findings underscore the importance of considering examiner effects in the broader context of banking supervision and regulation.

7 Conclusion

This paper provides a comprehensive analysis of the role of individual bank examiners in shaping supervisory outcomes and influencing future bank behavior. Using a rich dataset

covering U.S. bank holding companies from 2005 to 2023, we develop a novel measure—the examiner-specific effect (ESE)—that captures the difference between actual supervisory ratings and those predicted from publicly available bank data. Our analysis reveals several key findings that contribute to our understanding of bank supervision and its implications for the financial system.

First, examiner-specific effects are strongly related to differences in information access. Examiners who are geographically closer to the banks they assess or conduct on-site reviews deviate more from model-predicted ratings, consistent with access to valuable private information. In contrast, ratings of publicly traded banks exhibit smaller deviations, indicating that broader market transparency narrows informational asymmetries between examiners and banks.

Second, we find that individual examiner characteristics—particularly relative stringency or leniency—play a persistent role in supervisory outcomes. Examiners who have historically been tougher in prior examinations tend to assign stricter ratings in subsequent reviews, even after accounting for bank fundamentals and macroeconomic conditions. This persistence underscores the discretionary component of supervision.

Third, supervisory tone varies systematically over time. Examiners are generally more lenient in the periods preceding crises and more stringent in their aftermath, reflecting cyclical shifts in supervisory risk tolerance. This temporal pattern has important implications for understanding the dynamics of financial stability and regulatory responsiveness.

Most importantly, examiner-specific effects have tangible real-world consequences. Banks receiving tougher-than-expected ratings reduce asset growth, strengthen capital positions, and curtail lending in the following year. These effects are asymmetric: stricter assessments have pronounced effects, whereas lenient ones have muted but potentially risk-enhancing consequences, as evidenced by higher nonperforming loans.

Together, these findings reveal that supervision is not merely a rule-based process but one shaped by human judgment and informational context. Despite a uniform regulatory framework, examiner heterogeneity introduces variation that can both enhance and complicate supervisory outcomes. While such discretion can improve risk identification by incorporating soft information, it may also generate uncertainty for banks and affect the consistency of oversight.

From a policy perspective, our results highlight a fundamental trade-off: tighter examinations promote safety and soundness but may constrain credit supply, particularly during economic downturns. Recognizing and managing this balance is essential for maintaining both financial stability and macroeconomic resilience. Future research could build on our approach by exploring how supervisory heterogeneity interacts with stress testing, macro-

prudential policy, and crisis management frameworks.

Overall, our findings underscore that examiners are not passive enforcers of regulation but active participants in the information production process that underpins modern banking supervision. Understanding how their judgment, information access, and risk tolerance evolve over time is crucial for designing a supervisory system that is both consistent and adaptive to changing financial conditions.

References

Agarwal, Sumit, David Lucca, Amit Seru, and Francesco Trebbi, 2014, Inconsistent regulators: Evidence from banking, *The Quarterly Journal of Economics* 129, 889–938.

Agarwal, Sumit, Bernardo C Morais, Amit Seru, and Kelly Shue, 2024, Noisy experts? discretion in regulation, Working paper, National Bureau of Economic Research.

Angelini, Paolo, Francesca Lotti, Andrea Passalacqua, and Giovanni Soggia, 2021, The real effects of bank supervision: evidence from on-site bank inspections, *SSRN Work. Pap* 3705558.

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

Berger, Allen N and Sally M Davies, 1998, The information content of bank examinations, *Journal of Financial Services Research* 14, 117–144.

Berger, Allen N, Thomas Kick, and Klaus Schaeck, 2014, Executive board composition and bank risk taking, *Journal of Corporate Finance* 28, 48–65.

Beyhaghi, Mehdi and Jeffrey R Gerlach, 2024, Discount window borrowing 2003-2019, *Available at SSRN 4919012* .

Bonfim, Diana, Geraldo Cerqueiro, Hans Degryse, and Steven Ongena, 2023, On-site inspecting zombie lending, *Management Science* 69, 2547–2567.

Clement, Michael B, 1999, Analyst forecast accuracy: Do ability, resources, and portfolio complexity matter?, *Journal of Accounting and Economics* 27, 285–303.

Cole, Shawn, Martin Kanz, and Leora Klapper, 2015, Incentivizing calculated risk-taking: Evidence from an experiment with commercial bank loan officers, *Journal of Finance* 70, 537–575.

Correia, Sergio, Stephan Luck, and Emil Verner, 2025, Supervising failing banks, *Available at SSRN* .

Costello, Anna M, Joao Granja, and Joseph Weber, 2019, Do strict regulators increase the transparency of banks?, *Journal of Accounting Research* 57, 603–637.

Dang, Tri Vi, Gary Gorton, Bengt Holmström, and Guillermo Ordonez, 2017, Banks as secret keepers, *American Economic Review* 107, 1005–1029.

DeYoung, Robert, Mark J Flannery, William W Lang, and Sorin M Sorescu, 2001, The information content of bank exam ratings and subordinated debt prices, *Journal of Money, Credit and Banking* 900–925.

DeYoung, Robert, Dennis Glennon, and Peter Nigro, 2008, Borrower–lender distance, credit scoring, and loan performance: Evidence from informational-opaque small business borrowers, *Journal of financial intermediation* 17, 113–143.

Federal Reserve, 2019, Attachment two: Sr 19-4 / ca 19-3: Rfi rating system, Supervision and Regulation Letter SR 19-4 / CA 19-3, Board of Governors of the Federal Reserve System, accessed: [Insert date of access].

Flannery, Mark J, Simon H Kwan, and Mahendrarajah Nimalendran, 2004, Market evidence on the opaqueness of banking firms' assets, *Journal of Financial Economics* 71, 419–460.

Francis, Jere R and Dechun Yu, 2009, Big 4 office size and audit quality, *The Accounting Review* 84, 1521–1552.

Gopalan, Yadav, 2022, The effects of ratings disclosure by bank regulators, *Journal of Accounting and Economics* 73, 101438.

Gopalan, Yadav, Ankit Kalda, and Asaf Manela, 2021, Hub-and-spoke regulation and bank leverage, *Review of Finance* 25, 1499–1545.

Gul, Ferdinand A, Bikki L Jaggi, and Gopal V Krishnan, 2007, Auditor independence: Evidence on the joint effects of auditor tenure and nonaudit fees, *Auditing: a journal of practice & theory* 26, 117–142.

Hausman, Jerry A, 1978, Specification tests in econometrics, *Econometrica: Journal of the econometric society* 1251–1271.

Hirtle, Beverly, 2016, Public disclosure and risk-adjusted performance at bank holding companies, *Economic Policy Review, Issue Aug* 151–173.

Iannotta, Giuliano, 2006, Testing for opaqueness in the european banking industry: evidence from bond credit ratings, *Journal of Financial Services Research* 30, 287–309.

Kandrac, John and Bernd Schlusche, 2021, The effect of bank supervision and examination on risk taking: Evidence from a natural experiment, *The Review of Financial Studies* 34, 3181–3212.

Ke, Bin and Yong Yu, 2006, The effect of issuing biased earnings forecasts on analysts' access to management and survival, *Journal of Accounting Research* 44, 965–999.

Krainer, John and Jose A Lopez, 2009, Do supervisory rating standards change over time?, *Economic Review-Federal Reserve Bank of San Francisco* 13.

Morgan, Donald P, 2002, Rating banks: Risk and uncertainty in an opaque industry, *American Economic Review* 92, 874–888.

Srivastav, Abhishek and Jens Hagendorff, 2016, Corporate governance and bank risk-taking, *Corporate Governance: An International Review* 24, 334–345.

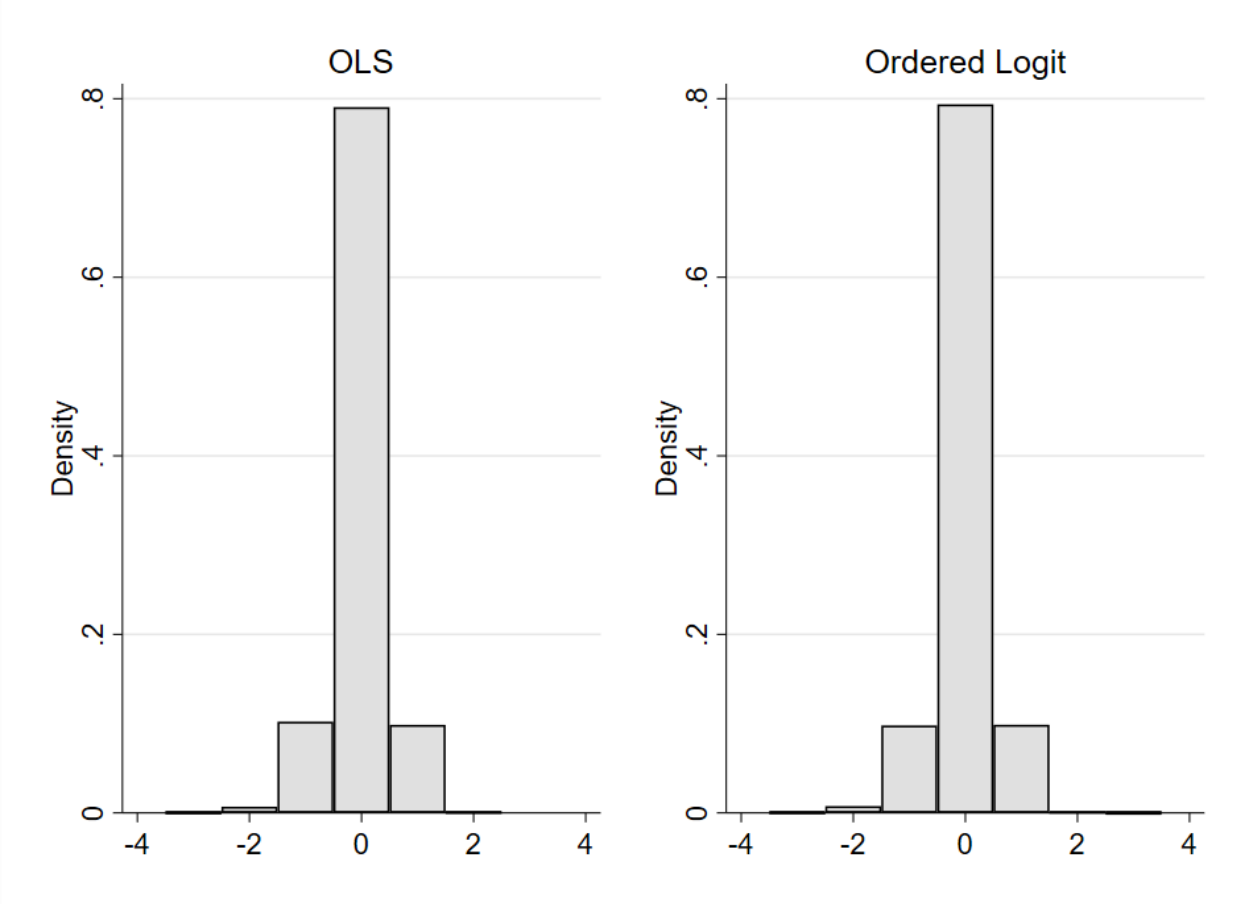


Figure 1: **Distribution of Examiner-Specific Effects.** This figure compares the distribution of estimated examiner-specific effects from two models: an ordinary least squares (OLS) model and an ordered logit (OLOGIT) model. The x-axis represents the difference between the predicted model rating and the actual examiner rating, with a value of zero indicating perfect agreement between the two. Positive values indicate that the examiner's rating was higher than the predicted model rating, while negative values indicate that the examiner's rating was lower.

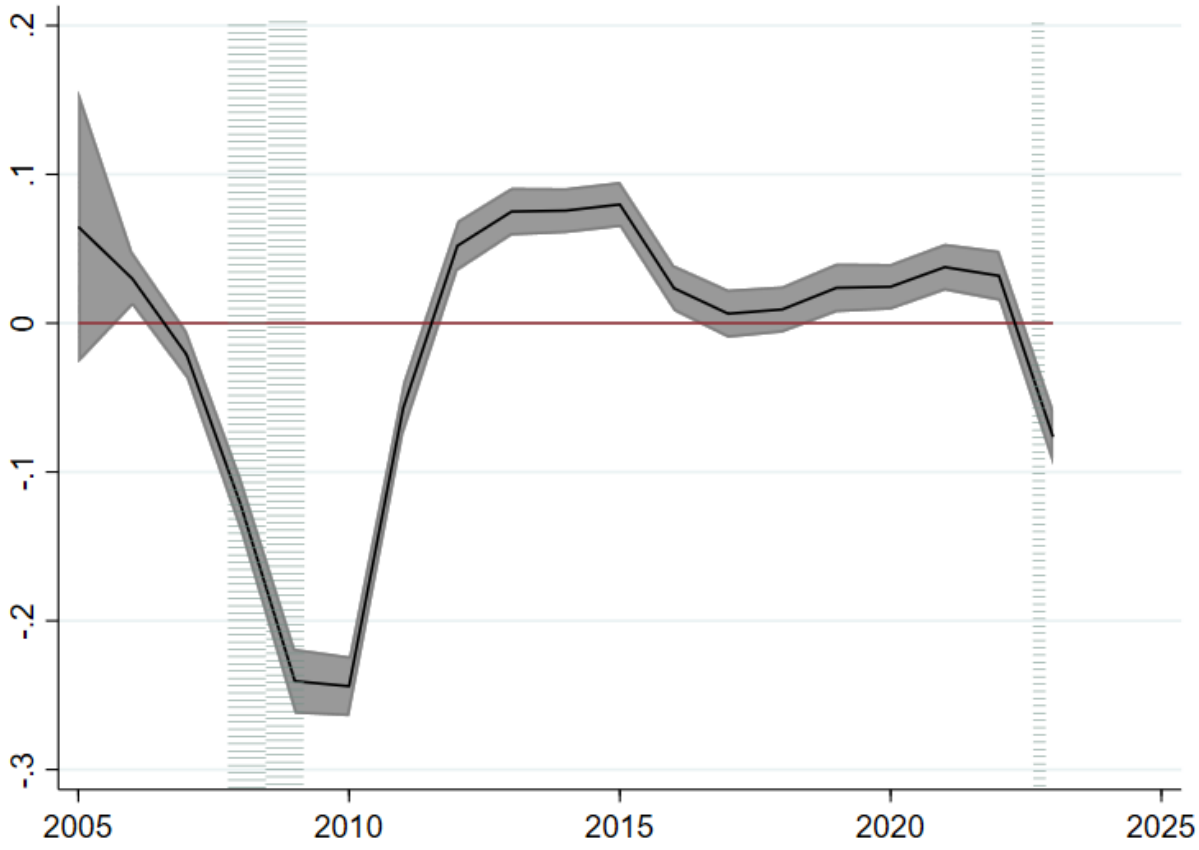


Figure 2: **Average Examiner-Specific Effect on Rating.** This figure presents the average examiner-specific effect on bank ratings, calculated as the difference between actual ratings and model-predicted ratings. The predictions use models that incorporate maximum available public information to forecast examination ratings. The figure plots the average examiner effect over time from 2005 to 2023, along with a 90% confidence band (5th-95th percentile). Shaded areas represent the Global Financial Crisis (2007Q4–2009Q2) and the banking turmoil of 2023Q1.

Table 1: **Distribution of RFI Composite Ratings**

This table presents the frequency distribution of RFI composite ratings. The sample include all bank holding companies that have filed FR Y9-C forms at least once during the sample during 2005–2023. Ratings range from 1 (strong) to 5 (unsatisfactory).

RFI Composite Rating	Freq.	Percent
1	5,957	23.60
2	14,474	57.33
3	3,070	12.16
4	1,095	4.34
5	649	2.57
Total	25,245	100.00

Table 2: **Summary Statistics**

This table presents summary statistics for a panel of 2,593 unique bank holding companies that filed Form Y-9C at least once between 2005 and 2023. The data are observed at the bank-year level. The variables used in this analysis are defined in the Appendix.

	N	Mean	SD	10%	Median	90%
Total Assets (thousand USD)	25,190	11,241,954	95,984,616	222,739	583,463	6,755,556
Log(Assets)	25,190	13.672	1.511	12.314	13.277	15.726
Equity/Assets	25,190	0.097	0.033	0.058	0.094	0.137
Securities/Assets	25,190	0.098	0.115	0.000	0.038	0.293
Loans/Assets	25,189	0.345	0.342	0.000	0.390	0.779
ROA	24,996	0.006	0.007	-0.001	0.007	0.015
Trading Assets/Assets	13,237	0.002	0.006	0.000	0.000	0.002
Allowance for Loans/Assets	13,237	0.010	0.005	0.005	0.009	0.018
Leverage Ratio	13,213	0.091	0.030	0.062	0.092	0.124
Double Leverage Ratio	13,237	0.005	0.018	0.000	0.000	0.005
Tier 1 Capital Ratio	12,707	0.131	0.042	0.087	0.123	0.186
Ratio of Residential Real Estate Loans	13,216	0.256	0.147	0.076	0.241	0.463
Ratio of Commercial Real Estate Loans	13,216	0.437	0.180	0.187	0.443	0.672
Ratio of Commercial and Industrial Loans	13,216	0.147	0.096	0.039	0.132	0.281
Ratio of Consumer Loans	13,216	0.053	0.068	0.003	0.027	0.146
Ratio of Depository Institution Loans	13,216	0.000	0.002	0.000	0.000	0.000
Ratio of Foreign Institution Loans	13,212	0.000	0.000	0.000	0.000	0.000
Ratio of Agriculture Loans	13,216	0.018	0.039	0.000	0.001	0.062
Loans Herfindahl-Hirschman Index	13,216	0.370	0.128	0.216	0.355	0.541
Assets Herfindahl-Hirschman Index	13,216	0.587	0.405	0.297	0.470	0.947
Risk-Weighted Assets/Assets	12,729	0.728	0.111	0.572	0.738	0.867
NPL/Assets	13,237	0.013	0.016	0.001	0.007	0.034
NPL/Loans	13,216	0.020	0.024	0.002	0.010	0.052
Loan Loss Reserve/Loans	13,216	0.016	0.008	0.008	0.013	0.027
Loan Loss Reserve/NPL	13,036	2.533	3.638	0.388	1.224	5.820
NPL-Residential Real Estate	13,117	0.018	0.022	0.001	0.009	0.046
NPL-Commercial Real Esate	13,119	0.024	0.033	0.000	0.010	0.071
NPL-Commercial and Industrial	12,789	0.015	0.020	0.000	0.007	0.040
NPL-Consumer	13,070	0.006	0.009	0.000	0.003	0.016
Market-to-Book Ratio	3,090	1.248	0.588	0.585	1.155	2.052
Monthly Stock Return (1 quarter) (%)	3,542	0.208	0.650	-0.483	0.098	1.000
Monthly Stock Return (1 year) (%)	3,500	0.072	0.292	-0.296	0.060	0.449
Stock Return Volatility	3,500	0.082	0.044	0.039	0.071	0.143
Average Stock Turnover	3,542	0.097	0.107	0.011	0.064	0.206

Table 3: **Benchmark Forecast Model**

This table presents the results of a benchmark forecast model that predicts examiner ratings based on bank fundamentals, using a set of ratios that are available for all banks. We estimate three alternative models with the same set of variables: an ordered logit model, a logistic model, and an ordinary least squares (OLS) model. T-statistics are shown below the parameter estimates in parentheses and are calculated using robust standard errors clustered by bank. *, **, and *** indicate statistical significance at the 10%, 5%, and 1% levels, respectively.

	Composite Rating		
	(1)	(2)	(3)
	Ordered Logit	Logit	OLS
Last Composite Rating	3.021***	2.769***	0.684***
	(78.053)	(50.452)	(103.869)
Log(Assets)	0.105***	0.121***	0.015***
	(6.976)	(4.724)	(4.932)
Equity/Assets	−8.322***	−9.303***	−1.884***
	(10.187)	(7.241)	(11.685)
Securities/Assets	−1.440***	−1.635***	−0.300***
	(5.700)	(4.251)	(5.969)
Loans/Assets	0.863***	0.859***	0.184***
	(9.605)	(6.760)	(9.937)
ROA	−136.416***	−181.396***	−32.211***
	(36.627)	(26.678)	(39.714)
Constant		−7.908***	0.820***
		(21.599)	(18.119)
Threshold for category 2 vs. 1	3.456***		
	(15.645)		
Threshold for category 3 vs. 1-2	8.635***		
	(36.992)		
Threshold for category 4 vs. 1-3	12.073***		
	(47.517)		
Threshold for category 5 vs. 1-4	15.233***		
	(54.585)		
Observations	22,754	22,754	22,754
Adj. R-squared			0.725
Pseudo R-squared	0.481	0.570	

Table 4: **Forecast Models Using All Available Public Information**

This table presents the results of three forecast models that predict the composite rating of a bank based on publicly available information. We report the adjusted R-squared for each model, which measures model fit and allows for comparison across the three models. The models differ in their complexity and the types of variables used: Model 1 (Baseline): Uses basic ratios obtained from financial statements, available for all banks, as described in Table 2. Model 2 (Saturated): Incorporates a wide range of ratios built from Y-9C filings. Model 3 (Market-Enhanced): Adds market-related information to the variables used in Model 2, including average monthly stock returns, market return volatility, turnover, and market-to-book ratio in the month preceding the examination. T-statistics are shown below the parameter estimates in parentheses and are calculated using robust standard errors clustered by bank. *, **, and *** indicate statistical significance at the 10%, 5%, and 1% levels, respectively.

	Composite Rating		
	(1) Baseline Model	(2) Saturated Model (from Statements)	(3) With Additional Market Variables
Number of Variables	6	29	34
Observations	3,419	2,707	2,684
Adj. R-squared	0.676	0.702	0.710

Table 5: **Relationship between Examiner-Specific Effect and Measures of Access to Private Information (1)**

This table examines the relationship between the absolute Examiner-Specific Effect on Rating, which measures the absolute difference between the model-predicted and actual rating, and various measures of distance between the bank and the examiner's office. Additionally, the table investigates the relationship between the absolute Examiner-Specific Effect on Rating and whether the exam was conducted online. A larger value of the dependent variable indicates that the model was less predictive of the actual rating based on available public information. T-statistics are shown below the parameter estimates in parentheses and are calculated using robust standard errors clustered by bank. *, **, and *** indicate statistical significance at the 10%, 5%, and 1% levels, respectively.

	Absolute Examiner-Specific Deviation							
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Distance Above Median	-0.021***		-0.020***				-0.022***	
	(2.773)		(2.676)				(3.150)	
Distance 25 to 50 Miles		-0.003		-0.005				-0.019
		(0.257)		(0.406)				(1.496)
Distance > 50 Miles		-0.033***		-0.031***				-0.044***
		(3.548)		(3.438)				(4.887)
Offsite Exam					-0.040***	-0.026**	-0.027***	-0.026**
					(3.713)	(2.360)	(2.599)	(2.525)
Examination Year FE			YES	YES		YES	YES	YES
Bank FE					YES	YES		
Controls							YES	YES
Observations	22,754	22,754	22,754	22,754	22,625	22,625	22,754	22,754
Adj. R-squared	0.001	0.001	0.027	0.027	0.077	0.097	0.040	0.041

Table 6: **Relationship between Examiner-Specific Effect and Measures of Access to Private Information (2)**

This table examines the relationship between the absolute Examiner-Specific Effect on Rating, which measures the absolute difference between the model-predicted and actual rating, and whether the bank is publicly traded. A larger value of the dependent variable indicates that the model was less predictive of the actual rating based on available public information. T-statistics are shown below the parameter estimates in parentheses and are calculated using robust standard errors clustered by bank. *, **, and *** indicate statistical significance at the 10%, 5%, and 1% levels, respectively.

	Absolute Examiner-Specific Deviation			
	(1)	(2)	(3)	(4)
Publicly Traded	-0.085^{***}	-0.074^{***}	-0.034^{***}	-0.042^{***}
	(10.508)	(9.173)	(3.673)	(3.817)
Examination Year FE		YES	YES	YES
Controls (Baseline)			YES	
Controls (Saturated)				YES
Observations	22,754	22,754	22,754	10,041
Adj. R-squared	0.005	0.030	0.039	0.052

Table 7: **Relationship between Examiner-Specific Effect and Examiner Stringency**

This table examines the relationship between the examiner-specific effect on a bank's rating and the examiner's stringency, as measured by their behavior in exams of other banks. We estimate several specifications that control for examination year and bank fixed effects, as well as baseline or saturated bank-level controls (excluding starting rating). The first five columns report the results of these regressions, with the examiner-specific effect as the dependent variable. In the last column, we use the starting rating as the dependent variable, to investigate whether examiners who are stringent in their evaluations of other banks are more likely to assign weaker ratings to banks. T-statistics are shown below the parameter estimates in parentheses and are calculated using robust standard errors clustered by bank. *, **, and *** indicate statistical significance at the 10%, 5%, and 1% levels, respectively.

	Examiner Specific Effect on Rating					Starting Rating
	(1)	(2)	(3)	(4)	(5)	(6)
Examiner Stringency	0.558***	0.380***	0.384***	0.376***	0.246***	-0.049
	(16.388)	(11.361)	(10.370)	(10.149)	(5.073)	(0.943)
Examination Year FE		YES	YES	YES	YES	YES
Bank FE			YES	YES	YES	YES
Controls (Baseline)				YES		
Controls (Saturated)					YES	YES
Observations	22,237	22,237	22,101	22,101	9,340	9,340
Adj. R-squared	0.019	0.048	0.049	0.054	0.048	0.686

Table 8: **Examiner-Specific Effects on Bank Ratings and Their Influence on Future Risk-Taking and Investments**

This table analyzes how examiner-specific effects on bank ratings correlate with future bank risk-taking behavior and investment decisions. All specifications include controls for examination year and bank fixed effects. We employ three measures of examiner-specific effects. The original measure is the difference between the actual rating and the rating predicted by a model using all available public information about the bank. Additionally, we introduce two new measures: the absolute value of the examiner fixed effect when the actual rating is harsher than the model-predicted rating (positive effect), and the absolute value of the examiner fixed effect when the actual rating is more lenient than the model-predicted rating (negative effect). T-statistics are shown below the parameter estimates in parentheses and are calculated using robust standard errors clustered by bank. *, **, and *** indicate statistical significance at the 10%, 5%, and 1% levels, respectively.

	Year $t+1$										
	Log(Assets)			Loan/Assets		Equity/Assets		Tier1 Ratio		NPL/Assets	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)
ESE_t	-0.018***										
	(4.179)										
ESE_t^+		-0.059***		-0.021***		0.003***		0.004***		-0.000	
		(6.597)		(3.450)		(4.629)		(3.519)		(0.502)	
ESE_t^-			0.010		0.005		-0.001		0.001		0.001***
			(1.608)		(1.398)		(1.522)		(1.044)		(3.454)
Examination Year FE	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES
Bank FE	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES
Observations	20,206	15,603	19,751	15,602	19,750	15,603	19,751	7,503	8,828	7,686	9,067
Adj. R-squared	0.972	0.975	0.975	0.734	0.723	0.769	0.766	0.756	0.752	0.694	0.669

Table 9: **Examiner Access to Private Information and Their Influence on Future Risk-Taking and Investments**

This table analyzes how the part of examiner-specific effects that is more likely related to the access to private information is related with future bank risk-taking behavior and investment decisions. All specifications include controls for examination year, bank, and examiner fixed effects. We employ three measures of examiner-specific effects. The original measure is the difference between the actual rating and the rating predicted by a model using all available public information about the bank. Additionally, we introduce two new measures: the absolute value of the examiner fixed effect when the actual rating is harsher than the model-predicted rating (positive effect), and the absolute value of the examiner fixed effect when the actual rating is more lenient than the model-predicted rating (negative effect). T-statistics are shown below the parameter estimates in parentheses and are calculated using robust standard errors clustered by bank. *, **, and *** indicate statistical significance at the 10%, 5%, and 1% levels, respectively.

	Year $t+1$										
	Log(Assets)			Loan/Assets		Equity/Assets		Tier1 Ratio		NPL/Assets	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)
ESE_t	-0.017***										
	(4.329)										
ESE_t^+		-0.056***		-0.013**		0.003***		0.005***		0.000	
		(6.640)		(2.061)		(5.230)		(3.759)		(0.020)	
ESE_t^-			0.012**		0.002		-0.000		0.002		0.001**
			(2.003)		(0.415)		(0.908)		(1.548)		(2.326)
Examination Year FE	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES
Bank FE	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES
Examiner FE	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES
Observations	19,808	15,185	19,346	15,184	19,345	15,185	19,346	7,064	8,397	7,243	8,633
Adj. R-squared	0.974	0.976	0.977	0.745	0.732	0.781	0.777	0.771	0.766	0.714	0.694

Appendix

Table A1: Variable Definition

The table provides variable definitions and specifies the source of data and the formula to calculate each variable using FR Y-9C mnemonics.

Variable	Definition	Source	Formula using FR Mnemonics
Composite rating	RFI/C(D) overall composite assessment	Federal Reserve	
Total Assets	Total assets (or natural log of total assets)	FR Y-9C, FR Y-9SP	BHCK2170, BHSP8519
Equity/Assets	Ratio of equity capital to total assets	FR Y-9C, FR Y-9SP	BHCK3210/BHCK2170, BHSP3210/BHSP8519
Securities/Assets	Ratio of total securities to total assets	FR Y-9C, FR Y-9SP	(BHCK1754 + BHCK1773)/BHCK2170, BHSP0390/BHSP8519
Loans/Assets	Ratio of total loans to total assets	FR Y-9C, FR Y-9SP	BHSP2122/BHCK2170, BHSP0390/BHSP8519
ROA	Ratio of net income to total assets (annual)	FR Y-9C	4×BHCK4340_qtr/BHCK2170, BHSP4340/BHSP8519
Trading Assets/Assets	Ratio of trading assets to total assets	FR Y-9C	BHCK3545/BHCK2170
Allowance for Loans/Assets	Ratio of allowance for loans to total assets	FR Y-9C	BHCK3123/BHCK2170
Leverage Ratio	Tier 1 leverage ratio	FR Y-9C	BHCK7204 (-2014) is replaced with BHCA7204 (2015-now)
Double Leverage Ratio	Ratio of lead bank's equity to parent's equity	FR Y-9C	BHCK3000/BHCK3210
Tier 1 Capital Ratio	Ratio of Tier 1 Capital to risk-weighted assets	FR Y-9C	BHCA8274/BHCAA223
Residential RE Loans	Ratio of residential RE loans to total loans	FR Y-9C	(BHDM5367 + BHDM5368 + BHDM1797)/BHCK2122
Commercial RE Loans	Ratio of commercial RE loans to total loans	FR Y-9C	(BHCKF158+BHCKF159 + BHDM1460 + BHCKF160+BHCKF161)/BHCK2122
C&I Loans	Ratio of C&I loans to total loans	FR Y-9C	(BHCK1763 + BHCK1764)/BHCK2122
Consumer Loans	Ratio of consumer loans to total loans	FR Y-9C	(BHCKB538+BHCKB539 + BHCKK137+BHCKK207)/BHCK2122
Depository Inst. Loans	Ratio of loans to foreign govts and official institutions to total loans	FR Y-9C	(BHCK2081)/BHCK2122
Agriculture Loans	Ratio of agricultural production loans to total loans	FR Y-9C	(BHCK1590)/BHCK2122
Foreign Inst. Loans	Ratio of loans to depository institutions to total loans	FR Y-9C	(BHCK1292+BHCK1296)/BHCK2122

Loans HHI	Herfindahl-Hirschman Index based on loan shares	FR Y-9C	(BHCKB538+4BHCKB539/BHCK2122)^2 + ((BHDM5367 + BHDM5368 + BHDM1797)/BHCK2122)^2 + ((BHCKF158_BHCKF159 + BHDM1460 + BHCKF160_BHCKF161)/BHCK2122)^2 + ((BHCK1763 + BHCK1764)/BHCK2122)^2
Assets Herfindahl-Hirschman Index	Herfindahl-Hirschman Index based on asset shares for credit card loans, residential real estate loans, commercial real estate loans, commercial and industrial loans, investment securities, and trading assets	FR Y-9C	(BHCKB538_4BHCKB539/BHCK2170)^2 + ((BHDM5367 + BHDM5368 + BHDM1797)/BHCK2170)^2 + ((BHCKF158_BHCKF159 + BHDM1460 + BHCKF160_BHCKF161)/BHCK2170)^2 + ((BHCK1763 + BHCK1764)/BHCK2170)^2 + ((BHCK1754 + BHCK1773)/BHCK2170)^2 + (BHCK3545/BHCK2170)^2
Risk-Weighted Assets/Assets	Ratio of risk-weighted assets to total assets	FR Y-9C	BHCAA223/BHCK2170
NPL/Assets	Ratio of nonperforming loans to total assets	FR Y-9C	BHCKA223/BHCK2170. Note: BHCKA223 (-2014) is replaced with BHCAA223 (2015-now).
NPL/Loans	Ratio of nonperforming loans to loans	FR Y-9C	(BHCK1407 + BHCK1403 + BHCK3506 + BHCK3507)/BHCK2122. Note: BHCK5525 (1990-2017) is replaced with BHCK1407 (-1990 & 2018-now), and BHCK5526 (1990-2017) is replaced with BHCK1403 (-1990 & 2018-now)
Loan Loss Reserve/Loans	Ratio of loan loss reserves to total loans	FR Y-9C	BHCK3123/BHCK2122
Loan Loss Reserve/NPL	Ratio of loan loss reserves to total loans	FR Y-9C	BHCK3123/BHCK2122. Note: BHCK5525 (1990-2017) is replaced with BHCK1407 (-1990 & 2018-now), and BHCK5526 (1990-2017) is replaced with BHCK1403 (-1990 & 2018-now)
NPL-Residential Real Estate	Ratio of nonperforming residential real estate loans to total residential real estate loans. Exclude BHCs where the category is ¡1% of total loans.	FR Y-9C	(BHCK5399 + BHCKC237 + BHCKC239 + BHCK5400 + BHCKC229 + BHCKC230)/(BHDM5367 + BHDM5368 + BHDM1797)
NPL-Commercial Real Estate	Ratio of nonperforming commercial real estate loans to total commercial real estate loans. Exclude BHCs where the category is ¡1% of total loans.	FR Y-9C	(BHCKF174+BHCKF175 + BHCKF176+BHCKF177 + BHCK3500 + BHCK3501 + BHCKF180 + BHCKF181 + BHCKF182 + BHCKF183)/(BHCKF158 + BHCKF159 + BHDM1460 + BHCKF160 + BHCKF161)

NPL-Commercial and Industrial	Ratio of nonperforming commercial and industrial loans to total commercial and industrial loans. Exclude BHCs where the category is ¡1% of total loans.	FR Y-9C	(BHCK1607 + BHCK1608)/(BHCK1763 + BHCK1764)
NPL-Consumer	Ratio of nonperforming commercial and industrial loans to total commercial and industrial loans. Exclude BHCs where the category is ¡1% of total loans.	FR Y-9C	(BHCKB576 + BHCKK214 + BHCKK217 + BHCKB577 + BHCKK215 + BHCKK218)/(BHCKB538 + BHCKB539 + BHCKK137 + BHCKK207). Note: BHCK5384 (-2000) is replaced with BHCKB576 (2001-now), and BHCK5385 (-2000) is replaced with BHCKB577 (2001-now)
Monthly Stock Return (1 year) (%)	Average monthly returns over 12 months	CRSP	
Stock Return Volatility	Std Dev of monthly returns over 12 months	CRSP	
Average Stock Turnover	Average turnover (volume/shares outstanding) over 12 months	CRSP	
Market-to-Book Ratio	Ratio of market capital to book equity	CRSP, FR Y-9C	(PRC×SHROUT)/BHCK3210
Monthly Stock Return (1 quarter) (%)	Average monthly returns over 3 months	CRSP	