Disciplining Banks through Disclosure: Evidence from CFPB Consumer Complaints*

Jeffrey Jou[1], Anya Kleymenova[2], Andrea Passalacqua[3], László Sándor[4], Rajesh Vijayaraghavan[5]

January 13, 2025

Abstract

We study the depositors' reaction to the disclosure of consumer complaints about their banks. Using the Consumer Financial Protection Bureau (CFPB) complaints data, we find banks subject to prudential CFPB oversight that receive consumer complaints experience a decline in uninsured deposits and respond by increasing offered deposit rates. We also analyze the content of consumer complaints in several ways. First, we find that complaints relating to bank accounts see a larger decline in deposits. Next, we leverage artificial intelligence (AI) tools to explore additional complaint classifications. Using topics identified by ChatGPT, we find complaints containing information on resolution expectations do not lead to similar deposit declines. Overall, our findings provide new evidence on the role of consumer complaints disclosure as a disciplinary mechanism and underscore the potential of AI tools to enhance the classification of complaints and support regulatory oversight.

JEL Classification: G21, G28, G53
Keywords: Disclosure, CFPB, Consumer complaints, AI, LLM, ChatGPT

*This paper was previously circulated as "Regulatory Disclosure and Access to Credit." We thank Anne Beatty (discussant), Orvin "Orv" Kimbrough (discussant), Lewis Kirvan, John Krainer, Allison Nicoletti (discussant), F. Dimas Pena-Romera (discussant), Rimmy E. Tomy (discussant), James Wang, Teng Wang, and seminar participants at the Federal Reserve Board, the AEA Annual Meeting, the JAAF-ATP India Symposium, Yale University, the Haskayne and Fox Accounting Conference, the IMF, the AAA Annual Meeting, the FDIC Bank Research Conference, and the Community Banking Research Conference for their helpful comments and suggestions. We are grateful to Arturo Bardales, Millennium Li, and Lori Leu for their excellent research assistance. The views expressed in this study are those of the authors and do not necessarily reflect the views of the Federal Reserve Board, the Federal Reserve System, or the Consumer Financial Protection Bureau. This article is the result of Andrea Passalacqua's independent research.

[1]The Wharton School, University of Pennsylvania, 3620 Locust Walk, Philadelphia, PA 19104; jouj@wharton.upenn.edu

[2]Federal Reserve Board, 20th Street and Constitution Avenue NW, Washington, DC 20551; anya.kleymenova@frb.gov

[3]Georgetown University McDonough School of Business, 37th and O Streets, NW, Washington DC 20057; andrea.passalacqua@georgetown.edu

[4]Consumer Financial Protection Bureau, 1700 G Street NW, Washington, DC, 20552; laszlo.sandor@cfpb.gov

[5]University of British Columbia, 2053 Main Mall, Vancouver, BC, Canada V6T 1Z2; rajesh.vijayaraghavan@sauder.ubc.ca

1. Introduction

The disclosure of regulatory information about firms could provide a potential disciplinary mechanism, influencing firm behavior. In the context of financial institutions specifically, such market discipline through disclosure might affect deposits and bank lending. Prior studies document that disclosure could increase monitoring by funding providers and improve bank operations (Diamond & Dybvig, 1983; Goldstein & Sapra, 2014; Anbil, 2018; Passalacqua et al., 2019; Kleymenova & Tomy, 2022; Granja & Leuz, 2024). However, a critical gap exists in the literature regarding the impact of disclosing *consumer complaints* as a form of market discipline. While regulatory disclosures are routinely scrutinized, the potential for consumer complaints to act as a signal of underlying firm issues and their subsequent impact on firm behavior remains less explored. We address this gap by investigating whether the public disclosure of consumer complaints, specifically through the Consumer Financial Protection Bureau (CFPB) database, provides a tool for depositors to exert market discipline on banks.

Our central research question is: Does the public disclosure of consumer complaints affect the market discipline exerted on banks, as evidenced by changes in depositors' behavior? To answer this question, we analyze the CFPB's consumer complaint database, which contains unstructured text, and evaluate depositors' reaction when complaints are disclosed. In our analyses, we utilize both traditional textual analysis methods and cutting-edge Artificial Intelligence (AI) tools, including Large Language Models (LLMs) such as ChatGPT, to extract meaningful insights from the unstructured complaint narratives.

Our analysis focuses on depositor reactions, as prior studies document that depositors, especially uninsured depositors, impose market discipline on banks by withdrawing their deposits in response to information disclosure. The large bank failures in March 2023 in the United States provide a stark reminder of the rapidity with which depositors can act, potentially leading to a bank's collapse. The failure of Silicon Valley Bank (SVB) serves as a compelling real-world example to underscore the central arguments and motivations behind the focus of our paper: the critical role of market discipline imposed by uninsured depositors. This case highlights the potential of consumer complaints to serve as early warning signals of underlying issues within a financial institution. SVB's failure demonstrates the consequences of failing to adequately address early warning signs. While SVB had some regulatory oversight, the escalation of concerns was not adequately reflected in traditional regulatory data until it was too late. Consumer complaints, however, *could* offer an earlier, more granular, and proactive indication of potential issues with a bank, equipping prudential and consumer regulators with additional tools for effective regulatory oversight.

Prior studies have examined depositors' reactions to negative information released by regulators (Diamond & Dybvig, 1983; Anbil, 2018; Kleymenova, 2018; Chen et al., 2022), but the effect of consumer complaints—a more direct reflection of customer experience—remains less understood. Our work builds upon the existing literature on market discipline and bank supervision and enforcement (An et al., 2024; Eisenbach et al., 2022; Granja & Leuz, 2024; Kandrac & Schlusche, 2021; Hirtle et al., 2020; Passalacqua et al., 2019; Eisenbach et al., 2017; Agarwal et al., 2014) by exploring the unique role of consumer complaints as a source of information about bank performance that could trigger a depositors' reaction.

Our analysis leverages the CFPB database of consumer complaints and links it to deposits. The CFPB was created as part of the Dodd-Frank Wall Street Reform and Consumer Protection Act of 2010 to provide oversight of financial consumer markets. It began operations in 2011 and has supervisory authority over banks and nonbanks in three main areas: rule-making, supervision and examination, and enforcement. Since 2012, the CFPB has published consumer complaints about financial institutions that fall under its supervision. Depository institutions and their affiliates with total assets above $10 billion and all financial services providers with retail consumers fall under the CFPB's oversight. For firms that the CFPB oversees, it publishes consumer complaints after the affected firm responds, confirming a commercial relationship with the consumer or after 15 days of receiving a complaint, whichever comes first. The complaints provide details about the corresponding financial product and contain unstructured text in the comments received from consumers.[1] For institutions that do not fall under its supervision, CFPB forwards consumer complaints to the corresponding regulator and does not publish them in the database.[2]

We focus on commercial banks with total assets between $1 billion and $25 billion to avoid the confounding effects of different regulations applied to larger banks and to be close to the $10 billion threshold for CFPB supervision (Fuster et al., 2021). Our sample includes over 700 banks, with approximately 77 supervised by the CFPB at various points during the sample period. The data allows for a granular analysis of complaints across several financial products (e.g., mortgages, credit cards, and bank accounts). A major strength of our dataset is the inclusion of consumer-submitted narratives from June 2015. This allows a deeper understanding of the nature of complaints that go beyond standard categorizations.

Our empirical strategy employs a range of econometric techniques, starting with descriptive analyses to establish the materiality of complaint disclosures. We then assess the impact of disclosed complaints on several key outcome variables, including total deposits, insured deposits, and uninsured deposits. Additionally, We investigate whether banks respond to declines in deposits by increasing offered deposit rates. In our analyses, we consider several measures of consumer complaints, including total complaints, account-related complaints, and the proportion of complaints containing narratives. By utilizing a continuous variable for complaints, we provide a more nuanced examination of the relationship between complaint disclosure and market discipline outcomes, advancing beyond prior studies that primarily focused on the binary presence or absence of complaints.

A novel aspect of our analysis lies in the use of AI tools (specifically ChatGPT) to deepen our understanding of the qualitative aspects of consumer complaints. We employed Chat-

[1] Unstructured text refers to the consumers' ability to write their own narratives, without guidance, about the issue(s) they are facing. Other aspects of the complaint form submitted to the CFPB include standardized options for the consumer. For example, consumers must choose from a set list of options for the product or service (e.g., mortgage, student loan, credit card, etc.) that matches their complaint.

[2] For example, complaints about commercial banks that are below the $10 billion size threshold for CFPB oversight are referred to their corresponding banking regulator (e.g., the Office of the Comptroller of the Currency (OCC) or the Federal Deposit Insurance Corporation (FDIC)).

GPT to generate several measures derived from the narratives, including those related to *Consumer Sophistication* (based on the assessment of the complexity and clarity of consumer narratives, considering factors such as regulatory awareness, financial terminology, and clarity) and *Resolution Expectations* (based on the level of detail and specificity in the consumers' demands for resolution). A unique aspect of our analysis is that ChatGPT was not merely employed as a computational tool but as a collaborator in generating tailored measures derived from the narratives. These measures, along with standard textual analysis techniques, allowed us to examine the nuances of the complaints and determine the degree to which they reflected negative sentiment and disappointment from consumers. By leveraging ChatGPT's contextual understanding, we uncovered dimensions of consumer behavior such as confidence in institutional accountability and the articulation of grievances that traditional methods alone might possibly overlook. This innovative approach underscores the potential of AI tools to enhance the study of qualitative data in regulatory and financial contexts.[3]

Our regression results demonstrate that public disclosure of consumer complaints is positively associated with market discipline. In particular, we find that banks receiving complaints see a decline in deposits, especially uninsured deposits. We find stronger results for complaints that are specifically focused on consumer accounts. Interestingly, we do not observe a stronger association between complaints disclosure and a decline in deposits for complaints with a higher consumer sophistication score, suggesting no differential response to those types of complaints from either insured or uninsured depositors. Consumer complaints could also indicate whether consumers expect their complaints to be resolved. We find that complaints with higher resolution expectations, see a higher inflow of uninsured deposits.

One potential concern with our findings could be that our measures merely capture

[3]The specific methodology behind generating and validating these measures is detailed in Appendix OA, which includes iterative refinement using LLM feedback.

poor bank performance. To address this concern, we evaluate whether banks with poorer performance (proxied by their return-on-assets, ROA) are more likely to see a decline in deposits when consumer complaints are published. We do not find a significant difference between depositors' response, indicative that our findings are not driven by bank performance per se.

Finally, we investigate whether banks respond to a decline in deposits by increasing deposit interest rates, particularly for longer-term deposits, to mitigate deposit outflows. We find that banks receiving a large number of complaints are more likely to increase their deposit rates for longer-term deposits, consistent with banks trying to preserve their deposit franchise when deposits decline.

Our findings provide compelling evidence of a new channel for market discipline on banks, driven by the public disclosure of consumer complaints. The results demonstrate that consumer complaints are not merely a measure of customer dissatisfaction; they serve as a significant signal of potential financial distress for banks, particularly for those under regulatory oversight. The decline in deposits, particularly uninsured deposits, suggests that depositors are actively monitoring bank performance and responding to negative information. The increase in deposit rates is a further indication of banks attempting to manage the fallout from negative publicity generated by the complaints.

The use of AI in this study highlights its potential to generate valuable insights from unstructured data, enriching our understanding of market behavior and providing novel methodologies to researchers in both accounting and finance. Our findings hold important implications for regulatory oversight and policymaking. In particular, our findings suggest that AI tools might help analyze large amounts of unstructured texts, such as consumer complaints and might be useful for identifying or classifying issues. Recent announcements by US Federal Banking Regulators highlight that AI tools are being utilized for analyzing large amounts of unstructured data to improve operational efficiency.[4]

[4] See for example, the summary in Risk.net's June 27, 2024 article available at

Our paper contributes to several strands of the economics, finance, and accounting literature. First, we contribute to the growing literature on consumer financial protection. For instance, Fuster et al. (2021) study the effect of the introduction of the CFPB on mortgage lending by taking advantage of the size threshold employed to identify which banks are supervised by the CFPB. The authors find that the creation of the CFPB reduces supervised bank lending to riskier borrowers and reduces foreclosures. Our work focuses only on banks under CFPB supervision and shows that consumer complaints to the CFPB and the disclosure of these complaints provide a new channel for market discipline in addition to increased regulatory oversight. Hayes et al. (2021) study the different patterns of consumer complaints according to different levels of consumer trust in financial institutions. They distinguish between low and high social trust areas and show that consumers in low-trust areas are less likely to trust banks, be more informed about the current regulations (and the potential violation of a law by a bank), and more likely to submit a formal complaint to the CFPB. As a result, banks are more likely to cut fees in counties with low levels of trust. In our study, we focus on the effect of public disclosure of complaints on all banks' consumers and in all geographic areas where they operate.

Second, several research papers study the role of consumer complaints disclosure. Our paper is closely related to Dou & Roh (2023) and Mazur (2022). Dou & Roh (2023) study the impact of consumer complaints disclosure on mortgage applications, while Mazur (2022) investigates the effect of complaints disclosure on mortgage application approval. We depart from their work by focusing on the disciplining role of complaints disclosure and depositors' reaction. We also classify and investigate the role of disclosure content by concentrating specifically on complaints narratives and not only instances of consumer complaints. To the best of our knowledge, we are the first to investigate the contents of these narratives.

More generally, our study relates to the literature investigating the effects of bank super-

https://www.risk.net/risk-management/7959632/us-fed-reveals-its-five-use-cases-for-generative-ai and the November 22, 2024 speech by Governor Michelle Bowman, available at https://www.federalreserve.gov/newsevents/speech/bowman20241122a.htm.

vision and enforcement on bank behavior and outcomes (An et al., 2024; Eisenbach et al., 2022; Granja & Leuz, 2024; Kandrac & Schlusche, 2021; Hirtle et al., 2020; Passalacqua et al., 2019; Eisenbach et al., 2017; Agarwal et al., 2014). While this line of work focuses more on the role of bank supervision and its implication on risk, credit supply to non-financial firms, and financial stability, we focus our attention on the supervision related to consumer financial protection and depositors' role in imposing market discipline.

Finally, our paper also contributes to the growing body of literature that leverages AI tools to enhance the analysis and classification of unstructured data. For example, Kim et al. (2023) utilize LLMs to summarize complex corporate disclosures and find that generative-AI-based summaries contain helpful information for investors. Kim et al. (2024b) and Kim et al. (2024a) show how AI can help uncover additional information for investors. Bybee (2023) constructs economic sentiment using large language models and historical news. Unlike these studies, we focus on the potential usefulness of AI tools for prudential regulation and supervision in the banking sector.

The rest of the paper is organized as follows. Section 2 describes the institutional setting. Section 3 describes the data and the construction of textual analysis measures using standard and AI-suggested techniques. Section 4 discusses the empirical model and presents our findings. Finally, section 5 concludes.

2. Institutional Setting

The CFPB is a federal agency created in 2010 as part of the Dodd-Frank Wall Street Reform and Consumer Protection Act. Its primary objective is to protect consumers in the financial marketplace by regulating and supervising financial institutions and enforcing consumer protection laws. In pursuit of this mission, one of CFPB's main functions is to operate and maintain the consumer complaints database, which allows consumers to submit complaints related to financial products and services. The CFPB reviews these complaints and takes appropriate actions, including coordinating with companies to address consumer issues.

The CFPB's consumer complaints process is designed to be user-friendly and accessible to all consumers. Complaints can be submitted via the CFPB's website, phone, or mail. Once a complaint is in the computerized system, it is routed to the financial institution to seek a resolution. Complaints about institutions that the CFPB does not supervise are routed to other agencies, as relevant. Additionally, the CFPB uses aggregate consumer complaints data to identify trends and patterns in the financial marketplace, which can help inform policy decisions and enforcement actions.

The financial institution subject to a consumer complaint is expected to respond to the complaint within 15 days and to provide a substantive response within 60 days. The time frame for when a complaint is published online can vary depending on the time it takes for the CFPB to review and process the complaint. Once the complaint is processed, it is published on the CFPB's website.[5] The online database of consumer complaints can be searched by company name, product, and issue to help consumers make more informed decisions. Generally, all complaints routed to companies will be published after 15 days; if the company responds earlier than 15 days, they will be published the next day. 98% of complaints sent to companies get timely responses. The CFPB removes personal information, such as the complainant's name and address, before publishing complaints.

Over time, the CFPB consumer complaints database has undergone substantive changes. Figure 1 presents a visual timeline of the dates of these changes. Soon after the enactment of the Dodd-Frank Act and the establishment of the CFPB in 2011, the CFPB began accepting complaints on July 21, 2011. Initially, the CFPB only accepted complaints related to credit cards. On June 19, 2012, the CFPB published its initial wave of complaints, populated with all credit card complaints received on and after June 1, 2012. In October 2012, the database was expanded to include consumer credit card complaints dating back to December 1, 2011.

[5] For a detailed description of the process see annual reports to Congress: CFPB Consumer Response Annual Report 2024.

Since its inception in 2011, the CFPB has adopted a phased-in approach to expand the types of complaints it accepts. As of 2023, the CFPB accepted consumer complaints across 11 categories of financial products, including, but not limited to, mortgages, debt collection, and virtual currencies.[6]

To promote the goal of providing consumers with timely and understandable information about consumer financial products and services, the CFPB began publishing consumer complaint narratives on June 25, 2015 for complaints dating back to March 19, 2015. Complaints published prior to this date did not allow consumers the opportunity to share their own experiences. Complaint narratives are published in the database on an opt-in basis. Consumers must opt in and allow the CFPB to publish their narratives. Published narratives are scrubbed of personal information such as bank account numbers, names, and addresses.

While the CFPB publishes the majority of complaints it receives, in some cases, the CFPB will not publish certain complaints. Complaints must meet all of the publication criteria in the CFPB's policy statement,[7] narrative data policy statement,[8] and narrative scrubbing standard.[9] For example, the CFPB may not publish a complaint for some irresolvable anomaly. Additionally, the CFPB may not publish a complaint if it is still under investigation and has not yet been closed. Consumers can close their complaints with or without a response or resolution. Consumers can also request the complaint not be published. If the consumer chooses to close the complaint or requests it not to be published, the CFPB will not post it in the public database.

The CFPB's supervisory authority applies to many financial institutions, such as banks and credit unions, as well as other non-bank financial institutions, such as credit bureaus

[6] Financial product types covered by the CFPB complaints database are: (1) checking or savings account; (2) credit card; (3) credit reporting or other personal consumer reports; (4) debt collection; (5) debt or credit management; (6) money transfer, virtual currency, or money service; (7) mortgage; (8) payday loan, title loan, personal loan, or advance loan; (9) prepaid card; (10) student loan; (11) vehicle loan or lease.

[7] Policy Statement

[8] Narrative Data Policy Statement

[9] Narrative Scrubbing Standard

and debt collectors. For banks, the CFPB only supervises institutions (and their affiliates) with more than $10 billion in assets and non-bank entities that are larger than participants in a market for a particular consumer financial product or service. Consumers can submit complaints even for companies the CFPB does not directly supervise. The CPFB collects these complaints and refers them to the appropriate supervisory agency. Complaints about banks are generally forwarded to the correct regulator in near real-time. The CFPB also securely shares complaint information with other federal, state, and local agencies to facilitate supervision and enforcement activities as well as monitor the market for consumer and financial products. In particular, through CFPB's secure Government Portal, local governments and other federal agencies have access to more granular information about consumers' complaints and companies' responses.[10]

As of March 2024, the CFPB has collected and published nearly 5 million complaints.[11] Credit reporting, credit repair services, or other personal consumer reports are the most common source of complaints and comprises more than half (3.1 million) of all complaints.[12] Many of these complaints are filed against the three largest credit bureaus: Equifax, Transunion, and Experian. Other common sources of complaints relate to debt collection, mortgages, and bank account services. We provide descriptive details for the complaints we study in the next section.

3. Data

This section describes the construction of our dataset, outlining the sample selection process, the use of ChatGPT to develop novel metrics for examining complaint narratives, and the descriptive characteristics of key variables. It also highlights the outcome measures

[10] For more details, see CFPB Blog.

[11] For our sample period spanning December 2011 to December 2019, nearly 1.5 million complaints have been published on CFPB's public database.

[12] This count is the sum of the following three distinct categories: "credit reporting, credit repair services, or other personal consumer reports," "credit reporting or other personal consumer reports," and "credit reporting."

used to examine depositor behavior and institutional responses.

3.1. Sample Selection

We focus our empirical analyses on commercial banks and savings banks, obtaining financial data from the quarterly Consolidated Reports of Condition ("Call Reports") collected by the Federal Financial Institutions Examination Council (FFIEC). Every national bank, state member bank, insured state nonmember bank, and savings association is required to file a Call Report as of the close of business on the last day of each calendar quarter. Call Reports contain data on bank income statements, balance sheets, and bank characteristics, such as institution name, operating city, state, and ZIP code.

We collect data beginning in 2010 Q1 through 2019 Q4, generating a ten-year panel dataset.[13] We follow a similar methodology laid out in Fuster et al. (2021) to construct our sample of banks. Specifically, we restrict our bank sample to commercial banks between $1 billion and $25 billion in total assets as of 2019 Q4 to limit the potential confounding effects of regulations imposed on larger banks. This threshold is chosen to focus on banks near the $10 billion threshold for CFPB supervision. After imposing these restrictions, our sample consists of 722 banks, 77 of which were supervised by the CFPB for at least one quarter within the sample period. Figure 2 displays the distribution of banks within our sample segmented by whether the bank was supervised by the CFPB or not. The total number of banks in our sample fluctuates mildly quarter-to-quarter, hovering around 700 banks on average. The number of CFPB-supervised banks in the sample increases over time, reflecting growth in bank assets. Initially, when the CFPB was established, only 20 banks were within the asset thresholds for our sample. The number of banks grows to 77 during the final quarter of our sample period.

The number of complaints a bank receives varies by bank size. Table 1 describes the

[13] We end our sample period in 2019 Q4 to avoid confounding effects from government interventions on financial institutions in response to the COVID-19 pandemic.

distribution of the complaints issued against financial institutions.[14] Assets are based on financial information reported on the 2019 Q4 Call Reports. While the CFPB supervises financial institutions with more than $10 billion in total assets, there are a number of smaller affiliate banks under $10 billion that also receive complaints. In general, the number of complaints increases with bank size. This result is consistent with a size effect, i.e., as larger banks have more depositors and borrowers and offer more financial services, they also receive more complaints, unconditionally. The skew is significant, with the majority of complaints filed against banks with more than $25 billion in total assets. Banks in our sample capture approximately 2% of total complaints issued against financial institutions across our sample period.

Focusing on the characteristics of banks in our sample, Table 2 provides summary statistics of key variables. Unsurprisingly, CFPB-supervised banks are larger and hold more deposits. Banks under CFPB supervision have an average of $11.1 billion in assets compared to $2.0 billion for non-CFPB-supervised banks. This difference in average bank size reflects the $10 billion asset threshold for CFPB supervision. Likewise, banks under CFPB supervision hold, on average, $8.3 billion in total deposits whereas non-CFPB supervised banks hold $1.6 billion in deposits. Other bank characteristics are approximately similar. CFPB-supervised banks are moderately more liquid, with average liquidity ratios 7% compared to an average of 6% for non-CFPB supervised banks. Moreover, banks supervised by the CFPB hold more capital (as measured by the Tier 1 capital ratio), likely due to greater regulatory scrutiny and safeguards. However, both groups of banks have capital ratios of at least 10%, far higher than the 6.5% regulatory minimum for well-capitalized banks. Furthermore, supervised banks are slightly more profitable as measured by their ROA.

[14]Financial institutions are defined as entities that appear in the FFIEC active or closed banks or the FDIC's Summary of Deposits database.

3.2. Consumer Complaints Data

The CFPB collects complaints about consumer financial products and services and sends them to companies it regulates for a response or forwards them to other prudential regulators. The CFPB consumer complaints data includes data from *certain* consumer complaints submitted on or after December 1, 2011.[15] Complaints in the database contain information about the product (e.g., savings account, credit card, or mortgage), the issue (e.g., managing an account or struggling to repay the student loan), the institution, and the geographic location of the complaint (e.g., ZIP code). Starting in June 2015, the CFPB began disclosing the consumer-submitted narratives from complaints received in March 2015 with the consumers' consent. The first publicly disclosed complaint with a narrative was on March 19, 2015. Within the portion of complaints filed against financial institutions between March 19, 2015 through the end of our sample period of December 31, 2019, 44.5% of complaints included a narrative.

We download the consumer complaints directly from the CFPB consumer complaints website.[16] Our sample period for the complaints data spans a decade beginning January 1, 2010, approximately two years prior to the initial inception of the complaints database in June 2012, through December 31, 2019. Approximately 1.5 million complaints were received and disclosed in the database within this period. From the consumer-reported "Company Name" field, we string-match the name to bank names reported on the Call Reports, keeping only the observations that have non-missing RSSD IDs, effectively filtering out complaints related to non-financial institutions (e.g., credit bureaus and debt collectors). This restriction reduces the number of complaints to 638,103 complaints and represents approximately 43% of all complaints. Within this subsample, 10,052 complaints were filed against banks within

[15]The CFPB only publishes complaints for institutions supervised by the CFPB. Complaints filed against firms outside the CFPB's jurisdiction are rerouted to the appropriate regulatory agency (e.g., Federal Reserve or OCC). In addition, the CFPB conducts a cursory review of the complaints and removes complaints deemed "unverifiable." These complaints are not published.

[16]See, https://www.consumerfinance.gov/data-research/consumer-complaints/#download-the-data

the \$1 and \$25 billion asset thresholds. Figure 3 shows that complaints matched to banks within our sample are geographically diverse, with complaints originating in more than 1,200 unique counties, representing more than one-third of the total counties in the United States. Unsurprisingly, complaints are concentrated in the most populous cities such as New York City, Chicago, and Los Angeles (Figure 4).

Figure 5 plots the quarterly number of complaints filed against all financial institutions identified within our sample period. The red line shows the best-fit line over time. Several factors contribute to the increase in the number of complaints over time. First, consumers may be more aware of the CFPB complaints database and, in turn, file more complaints. Indeed, there is a sharp uptick in the number of complaints over the first few quarters from 2011 Q4 to 2012 Q2 as awareness of the database spread. Second, the CFPB continued to add more product types under its purview over time. Initially, the CFPB only accepted complaints related to credit cards in 2011. Since then, the CFPB has expanded complaint product types to 11 different categories, drastically increasing the number of published complaints.

Past papers examining the CFPB complaints have focused on mortgage-related complaints and mortgage lending (Begley & Purnanandam, 2021; Dou & Roh, 2023; Dou et al., 2024; Mazur, 2022). Our study differs by incorporating all complaints. Figure 6 plots the total number of complaints over time, splitting out mortgage-related complaints from all other complaint types. While mortgages constitute an important aspect of banks' lending portfolios, the share of mortgage-related complaints relative to total complaints declines over time. This trend suggests incorporating other complaints outside of mortgages is important. Moreover, our study focuses on deposits. As such, we conduct additional tests focusing on bank account-related complaints exclusively. In contrast to the number of mortgage-related complaints, the share of account-related complaints has remained steady, with a slight increase, across the sample period.

3.3. ChatGPT and Natural Language Processing (NLP)

In addition to data provided within the consumer complaints database, we leverage the natural language processing (NLP) power of ChatGPT to analyze text contained in the consumer complaint narratives. Specifically, we prompted ChatGPT to generate novel measures based on ChatGPT's own assessment of the complaint narratives. With minimal user guidance, ChatGPT produced two measures – narrative sophistication and resolution expectancy, which were used in conjunction with our depositor reaction analysis.[17] We discuss the data generation process in the subsequent sections and provide a step-by-step discussion in Appendix OA.

3.3.1. Development of the Sophistication Measure

The sophistication measure assesses the complexity and quality of consumer complaint narratives submitted to the CFPB. The measure reflects varying levels of regulatory awareness, financial literacy, and communication skills, making them valuable for analyzing consumer behavior and institutional accountability. When prompting ChatGPT to generate this measure, we used two documents provided by the CFPB to contextualize the narratives: (i) CFPB Narrative Scrubbing Standards and (ii) CFPB Supervision and Examination Manual.[18] The scrubbing standards detail the types of information redacted from the complaints database while the supervision and examination manual outlines expectations for consumer protection and complaint resolution. After processing these initial documents, we provided ChatGPT a file containing consumer complaint narratives uploaded during the development process. This dataset allowed us to observe real-world examples of narratives and understand the variability in structure and content.

The sophistication measure evolved through an iterative prompting process. The initial

[17] In addition to these measures, ChatGPT also generated measures assessing the severity, sentiment, and escalation tendency based on text in the complaint narratives. These additional tests can be found in Appendix OA.

[18] See CFPB Narrative Scrubbing Standard and CFPB Supervision and Examination Manual

file containing a representative sample of consumer narratives, illustrated the presence of regulatory terms, financial jargon, and redactions. By examining this dataset, we identified variability in narrative quality, which informed the dimensions of sophistication. Next, based on the dataset and prompts, ChatGPT proposed dimensions such as regulatory awareness, financial terminology, and specificity. Narratives referencing specific laws (e.g., "Fair Credit Reporting Act") or containing terms like "interest rate" demonstrated higher sophistication. Furthermore, ChatGPT suggested incorporating Term Frequency-Inverse Document Frequency (TF-IDF) analysis to quantitatively determine the importance of terms in a narrative relative to the entire dataset, helping distinguish regulatory or financial terms (e.g., "escrow account" or "loan modification") from generic language.

We tested metrics on subsets of the narratives and provided feedback to refine scoring logic. For example, when clarity and specificity metrics underperformed on overly concise narratives, adjustments were made to consider tone and logical flow. In total, ChatGPT suggested seven factors for the final sophistication measure: regulatory awareness, financial terminology, specificity, clarity and structure, professionalism and politeness, vocabulary complexity, and impact of reductions. Each narrative was scored along these dimensions. The final scores were aggregated to produce the "sophistication index" used in our analysis.

3.3.2. Resolution Expectation Metric and Other Behavioral Measures

In addition to utilizing guided prompts, we also employed ChatGPT to autonomously generate measures to enhance our analysis. In this exercise, rather than prescribing a specific metric to evaluate (e.g., sophistication), we allowed ChatGPT to proactively propose a range of potential measures. These measures were designed to assess depositor reactions based on the dataset, regulatory framework, and our overarching objective of understanding consumer complaints.

ChatGPT identified three key metrics as particularly relevant for our analysis: the resolution expectation metric, behavioral indicators, and action orientation. The resolution expectation metric measures the clarity and specificity with which consumers articulate their desired resolutions, such as refunds or corrective actions. Behavioral indicators, in contrast, capture consumer tendencies such as escalation (eg., references to legal action or external authorities), frustration (eg., expressions of urgency or dissatisfaction), and resolution orientation (eg., cooperative or solution-focused language). Action orientation quantifies the actionable requests in the narratives (e.g., "resolve" or "refund") and captures the extent to which consumers explicitly demand solutions. Together, these metrics offer a multidimensional view of consumer reactions and complaint dynamics.

These measures were not only informed by the textual features of the narratives but also contextualized within the regulatory and institutional landscape outlined in the CFPB Examination Manual and related documents. The development of these metrics, including iterative refinements based on researcher feedback, is detailed further in Appendix OA. These measures were subsequently integrated with outcome variables to examine specific textual traits of consumer narratives, providing new insights into the behavior and expectations of depositors.

3.3.3. Descriptive Statistics of ChatGPT Measures

Table 3 provides descriptive statistics on the textual analysis measures generated by ChatGPT.[19] Complaints in our sample have an average sophistication index of 0.24 (out of 1). Furthermore, sample complaints have an average resolution expectation of 0.32 (out of 10). Despite these low means, there is a significant spread in these measures. Standard deviations for the sophistication index and resolution expectation are 0.18 and 0.65, respectively.

To provide more description and initial validation of the measures generated by ChatGPT, we plot the average measures by product type. Figure 7 displays the product averages for the sophistication index. The "Mortgages" product category has the highest sophistication index, with an average of over 0.3 (out of 1). The higher score indicates mortgage-related complaints are more detailed and involve more complex financial concepts, likely

[19] The table includes additional variables labeled "Action Orientation", "Escalation Tendency", and "Frustration Urgency." The effect of these measures on deposits is included in Appendix OA.

reflecting the nature of mortgages which involve significant financial commitments and require more knowledge to navigate. Conversely, prepaid cards have the lowest sophistication score. Prepaid cards, relative to mortgages, require lower financial knowledge leading to lower sophistication scores in these types of complaints.

Figure 8 presents the average resolution expectation measure by product type. Products like auto loans, personal loans, and mortgages typically have higher resolution expectation metrics. This trend indicates that consumers dealing with loans expect more from companies in resolving their complaints, possibly because of the financial and emotional stakes involved. Products related to bank accounts (e.g., checking and savings accounts) tend to show lower expectation metrics. This might reflect the perception that issues with these products are simpler or less critical than those involving loans, leading to lower consumer expectations for resolution.

3.4. Outcome Measures

For our primary analysis, we examine whether consumers discipline banks through deposit flows. We begin by looking at depositor reactions through deposit data from the Call reports. Specifically, we examine the effect of consumer complaints on total deposits, insured deposits, and uninsured deposits. Economic intuition suggests uninsured depositors are more likely to act as monitors since uninsured deposits may not be recovered in the event of a bank failure.

To investigate whether banks respond to declining deposits by changing their offered deposit rates, we utilize quarterly deposit interest rate data from the RateWatch Scholar compiled by S&P Global Market Intelligence. RateWatch contains deposit interest rate data for retail and business products from 2001 to 2020. These data span over 7,500 financial institutions (e.g., banks, credit unions, savings, and loan associations) and across 96,000 locations. Moreover, RateWatch provides deposit interest rate data at the product level (e.g., CD, interest checking accounts, saving accounts, or money markets) and sub-product level, including product term lengths and dollar tiers. This data allows us to estimate rates offered on various financial products at a granular level.

4. Empirical Model and Results

4.1. Do Consumers File Complaints?

Prior to empirically testing the depositors' reaction to complaints, we first descriptively analyze the salience of the complaints database for consumers. Specifically, we examine the change in complaint volume around specific bank events and scandals. If complaints reflect underlying issues at the bank, then complaints should increase around these event dates.

4.1.1. Wells Fargo

For the first case study, we examine complaint activity around the Wells Fargo fake accounts scandal. In September 2016, Wells Fargo was fined $185 million and later settled criminal and civil lawsuits for $3 billion for fraudulently opening customer accounts between 2002 and 2016. Bank employees set up sham customer accounts using fake email accounts without a customer's consent. Figure 9 plots the quarterly number of complaints against Wells Fargo over our sample period of complaints. While the average number of complaints hovers around 600, there is a distinct spike in the number of complaints filed during the shaded region between 2016 Q2 and 2016 Q4. This increase in the number of complaints corresponds with Wells Fargo's fake account scandal, suggesting consumer complaints may act as a lender monitoring channel.

4.1.2. Citibank

The second case study investigates credit card complaints filed against Citibank. In 2015, Citibank launched a new campaign to promote its Citigold card by offering customers 50,000 American Airlines reward miles (approximately $500 in value). However, many customers did not receive the bonus miles and complained to CFPB against Citibank. The first shaded region in Figure 10 shows a stark increase in the number of credit card-related complaints in the period following the Citigold card promotion. Moreover, consumer posts in different forums around that period directly mentioned and encouraged filing complaints in the CFPB database, likely driving more traffic and complaints to the database.[20]

More recently, Citibank has drawn attention for its response to the onset of the COVID-19 pandemic in early 2020. The second shaded region in Figure 10 shows a dramatic increase in the number of complaints between 2020 Q1 and 2020 Q3. During this period, Citibank accounted for nearly 37% of pandemic-related complaints about credit cards. Furthermore, the increase in complaints drew the attention of national media outlets, like CNN, which reported that complaints stemmed from several avenues ranging from inflexible late fees and interest charges to refusal to assist consumers experiencing financial hardship.[21]

4.1.3. Capital One

The final case study focuses on credit reporting-related complaints at Capital One. Figure 11 highlights two high-profile data breaches that revealed sensitive customer information: (1) Equifax data breach in 2017 and (2) Capital One data breach in 2019. In the first event, the personally identifying data of hundreds of millions of people was stolen in March 2017. This information included names, social security numbers, and credit card numbers. As the threat of identity theft and fraud increased, so did the number of complaints. In the year after the data breach, credit report-related complaints to Capital One, one of the largest credit card issuers, quadrupled in response. A similar pattern occurred following Capital One's own data breach in July 2019. In the aftermath of this cyberattack, more than 100 million customers were exposed. Correspondingly, consumer complaints filed against Capital One to CFPB increased substantially in the year following the incident.

Overall, these case studies underscore the importance of complaints as a tool for consumers to file grievances against their financial institutions. In each instance, the number of complaints related to scandals or deficiencies at the bank dramatically increases. While the

[20] See posts in PointsCentric and Miles per Day for more details.

[21] CNN reports more details about the content of the complaints in their article, *An alarming number of people are complaining about their Citi credit card accounts*

figures presented in this section are only descriptive, there is evidence of a market disciplining effect on banks through consumer complaints.

4.2. Do Depositors React to Consumer Complaints?

After establishing that complaints reflect bank news events, we investigate whether depositors react to the disclosure of complaints. Following our identification strategy, we focus on banks around the threshold of \$10 billion for CFPB supervision (Figure 2) and estimate the following OLS model:

$$Y_{b,t} = \alpha + \beta_1 Complaints_{b,t} + \gamma X_{b,t-1} + \theta_t + \eta_b + \varepsilon_{b,t}, \tag{1}$$

where b, t correspond to *bank* and *year-quarter*, respectively. $Y_{b,t}$ refers to deposits (in log levels). To separate potential differential effects between insured and uninsured depositors, our outcome measure includes total deposits, insured deposits, and uninsured deposits. We include a set of one-quarter lagged bank-level variables, $X_{b,t-1}$, to control for bank-specific characteristics. These variables include (i) bank size (measured as the natural logarithm of total assets), (ii) liquidity ratio (measured as the ratio of cash and cash equivalent to total assets), (iii) non-performing loans ratio (measured as the ratio of the sum of 90 or more day delinquent loans and nonaccrual loans to total assets), (iv) ROA (return on assets, measured as net income divided by total assets), and (v) Tier 1 capital ratio (measured as a ratio of Tier 1 capital to total assets). θ_t and η_b are time fixed effects and bank fixed effects, respectively, used to control for unobserved heterogeneity across banks and time. Robust standard errors are clustered by bank and year-quarter, which allows for correlation in error terms within a given bank and serial correlation across quarters. Appendix A provides further details on our variable definitions.

Our variable of interest is $Complaints_{b,t}$. This is a continuous variable that measures the total number of complaints bank b receives in year-quarter t. For robustness, we augment our complaints measure to capture the proportion of bank account-related complaints (i.e., complaints related to "Bank account or service" and "Checking or savings account") relative to total complaints. We posit bank account complaints are most related to depositors. In both specifications, our coefficient of interest is β_1. If depositors discipline banks through deposit flows, then we predict a negative β_1 coefficient. This finding suggests depositors remove deposits from banks with greater numbers of complaints.

Table 4 presents the results of this estimation. Column (1) presents the aggregate effect of complaints on total deposits (insured and uninsured). While the coefficient is negative, it is not statistically significant at conventional levels. Columns (2) and (3) decompose total deposits into insured and uninsured deposits. While there is no statistically significant effect of complaints on insured deposits, there is a significantly negative effect on uninsured deposits. This result is consistent with economic intuition suggesting uninsured depositors are stronger monitors of banks given they face higher risks of losing deposits in the event of a bank failure.

To further sharpen our findings, we narrow our focus to complaints related to deposit accounts. We posit that account-related complaints are most germane to depositors as they correspond to services depositors are likely to face. As a result, depositors may react more strongly to these complaints as compared to other complaint products such as debt collection or student loans. Table 5 presents the results under the same regression specification, but using a continuous measure of account-related complaints rather than total complaints. When using bank account complaints, there is a statistically significant negative effect on total deposits (column 1). Decomposing the deposits into insured and uninsured deposits reveals the negative effect is driven by uninsured deposits (column 3). There is no statistically significant change in insured deposits (column 2). Again, this finding is consistent with the intuition that uninsured depositors are stricter monitors as they stand to lose their deposits in times of economic stress. Overall, our findings suggest some evidence of depositor reaction in response to the disclosure of consumer complaints consistent with depositors disciplining banks in line with prior studies (Anbil, 2018; Kleymenova & Tomy, 2022; Diamond

& Dybvig, 1983).

A potential alternative explanation for our results is that underlying issues at the bank, rather than complaints, are driving deposit outflows. To disentangle these two effects, we conduct cross-sectional analyses, splitting the sample based on above- and below-median profitability, proxied by bank ROA. The intuition behind this test is banks with high ROA are "better performing" banks with fewer underlying issues compared to banks with low ROA. If deposit outflows are driven by underlying issues, then the outflows should be concentrated in banks with low profitability.

Table 6 presents the results of this cross-sectional test using total complaints, regardless of product type. With the exception of uninsured deposits for low ROA banks (column 6), there is no statistically significant effect of complaints on deposits. To test for differences between high and low ROA banks, we compute Wald statistics, testing for statistically significant differences in coefficients of the two bank types across each deposit category. The p-values for the difference between columns (1) and (4), (2) and (5), and (3) and (6) are 0.223, 0.416, and 0.371, respectively. Since the p-values are higher than conventional levels of significance, we conclude there is no evidence in support of the alternative explanation that underlying issues, proxied by bank profitability, are driving deposit outflows.

As a further test, we repeat the cross-sectional analysis using account-related complaints. Table 7 presents the findings. Again, we test for differences in coefficients across bank profitability groups. The p-values comparing columns (1) and (4), (2) and (5), and (3) and (6) are 0.823, 0.950, and 0.777, respectively. Using these statistical tests, we reach the same conclusion as before and interpret the results as evidence against the alternative explanation.

4.3. Do Complaint Narratives Matter?

Having examined the effect of complaint disclosure on deposits, we explore whether depositors react to the content contained within complaints. To do so, we exploit the CFPB's decision to disclose consumer complaint narratives beginning in June 2015. Specifically, we re-estimate Equation 1, using consumer narratives as the variable of interest rather than the total number of complaints. Formally stated, we estimate the following OLS specification:

$$Y_{b,t} = \alpha + \beta_1 Complaint\ Narratives_{b,t} + \gamma X_{b,t-1} + \theta_t + \eta_b + \varepsilon_{b,t}, \tag{2}$$

where $Complaint\ Narratives_{b,t}$ is the proportion of complaints that include consumer narratives relative to the total number of complaints received by bank b in quarter t. Control variables and fixed effects are consistent with the definitions in Equation 1.

Table 8 presents the results of this estimation. Column (1) shows complaint narratives have a significantly negative effect on total deposits. Similar to previous results, this decline in deposits is concentrated in uninsured deposits (column 3) rather than insured deposits (column 2). Columns (4)–(6) test the incremental effect of complaint narratives over total complaints by including the total number of complaints as a separate variable in the regression model. The negative effect of complaint narratives remains significant for both total deposits and uninsured deposits. These results suggest depositors react more to the content of the narratives rather than just the number of complaints.

Similar to previous tests, we repeat our analysis focusing on account-related complaints and narratives. Table 9 shows the results of the estimation using the proportion of account-related complaint narratives. Again, account-related complaint narratives are associated with a statistically significant deposit outflow for total deposits and uninsured deposits. There is no statistically significant change in insured deposits. In columns (4)–(6), we examine the incremental effect of account-related narratives over the proportion of account-related complaints. The effect of account-related narratives on deposits remains statistically significant for total deposits but loses its explanatory power for uninsured deposits under conventional significance levels.

4.4. Textual Analysis of Complaint Narratives

After examining the effect of disclosing complaint narratives, we turn our attention to analyzing the text of the complaints and its effect on deposits. Intuitively, these analyses examine whether the content of the narratives, as opposed to the complaint and narrative disclosure, drives depositor reactions. For this test, we employ the metrics developed by ChatGPT for sophistication and resolution expectation. The sophistication measure is a continuous measure from 0 to 1 that assesses narratives along seven key dimensions: regulatory awareness, financial terminology, specificity, clarity and structure, professionalism and politeness, vocabulary complexity, and redactions. The resolution expectation metric is a continuous variable from 0 to 10 that measures the clarity and specificity of resolution demands, including mentions of monetary amounts or timelines.

To operationalize this test, we follow a similar framework as before, estimating the following specification:

$$Y_{b,t} = \alpha + \beta_1 Text\ Measure_{b,t} + \gamma X_{b,t-1} + \theta_t + \eta_b + \varepsilon_{b,t}, \tag{3}$$

where the variable of interest is $Text\ Measure_{b,t}$ representing the average complaint sophistication index or resolution expectation score for bank b in time t. Control variables and fixed effects are consistent with Equation 1. Importantly, since we require banks with complaint narratives to assign these textual measures, we alter our sample to only include CFPB-supervised banks. Moreover, complaints only began having narratives in March 2015. In turn, these analyses use a sample of 79 banks from 2015 Q1 to 2019 Q4.

Table 10 presents the results of this estimation using the sophistication measure. Across all measures of deposits (total deposits, insured deposits, and uninsured deposits), there is no statistically significant effect of complaint sophistication on deposit flows. In other words, higher degrees of sophistication in the consumer narrative, as measured by regulatory awareness and financial terminology, do not lead to changes in deposits. This finding indicates a somewhat surprising conclusion: more sophisticated depositors (in terms of their financial understanding) are no more likely to withdraw their deposits than unsophisticated depositors. Ex ante, economic intuition would suggest sophisticated depositors are more stringent monitors of banks and would be more sensitive to bank issues. However, our results do not support this view.

Table 11 presents the results of the estimation using the resolution expectation measure. While there is no change in total deposits and insured deposits, there is a statistically significant increase in uninsured deposits. This result suggests complaints with higher likelihoods of resolution lead to more uninsured deposits. At first glance, this may be a surprising result. However, the result is consistent with intuition. Complaints with higher resolution expectations allow for immediate responses from banks. The swift resolution of complaints reflects positively on the bank, leading depositors to revise their beliefs about the bank upwards. In turn, depositors increase their deposits. This result is concentrated in uninsured deposits as they are the most sensitive type of depositors. Overall, textual analyses of the complaint narratives indicate there are differential effects based on the content of the complaint narratives.

4.5. Do Banks React to Consumer Complaints?

Given our observed deposit outflow associated with complaints, we next investigate whether banks respond by increasing offer rates on various deposit products in an attempt to attract more deposits. In particular, we are interested in understanding whether banks with consumer complaints provide higher rates on deposits, especially on deposit products that impose withdrawal restrictions on depositors. Utilizing data available through RateWatch, we estimate Equation 1 with the dependent variable being the natural logarithm of deposit rates offered on certificate of deposits (CD) contracts. To account for heterogeneity in bank deposit rates across different geographic locations, the unit of analysis for these tests is a bank-county-quarter observation. We estimate the following OLS specification:

$$Y_{b,c,t} = \alpha + \beta_1 Complaints_{b,c,t} + \gamma X_{b,t-1} + \mu Z_{c,t} + \theta_t + \eta_b + \phi_c + \varepsilon_{b,c,t}, \tag{4}$$

where the outcome variable is the natural logarithm of CD deposit rates for bank b in county c during quarter t. For our variable of interest $Complaints_{b,c,t}$ we use three separate variations of complaint measures to capture the effect of complaint disclosure (*Public disclosure*), complaint intensity (*High complaint*), and total complaints (*Total complaints*) on CD rates. Appendix A provides additional descriptions of these explanatory variables. In all specifications, we include bank and county controls as well as year-quarter, county, and bank-level fixed effects to control for unobserved heterogeneity.

Table 12 presents the results of this estimation. Column (1) shows that banks increase offered rates following the public disclosure of complaints. This increase appears to be mostly driven by banks with an above-median number of complaints (column 2). Column (3) suggests that a larger absolute number of complaints in a given quarter also provides some explanatory power and is positively and significantly associated with higher offered rates on deposits. This finding suggests banks actively seek to replenish their deposits by offering higher rates on deposits that can be locked in for a period of time, consistent with prior studies showing that banks increase rates on deposits when facing declines in their core-deposits (Acharya & Mora, 2015). Overall, banks respond to deposit outflows associated with complaints by increasing offered rates on their deposit products to attract more deposits and dampen the potentially negative effect of the public disclosure of consumer complaints.

4.6. Modeling Complaints Based on Topics

In our final set of analyses, we further utilize NLP techniques and infer the main topics contained in the corpus of the text of complaints using the Latent Dirichlet allocation (LDA) method. This method treats each document (in our case, a complaint) as a mixture of topics and each topic as a mixture of words. This approach allows documents to contain multiple topics in terms of content rather than being separated into discrete groups. Following this approach, we identify five main topics that appear throughout the complaints database. Figure 12 aggregates the content of the complaints into topic classes. Panel A presents topics based on the text of the public CFPB complaints database. The topics can be identified as (i) credit cards, (ii) bank services, (iii) bank payments, (iv) bank loans, and (v) bank accounts.[22]

As Figure 12, Panel A shows, the five topics are somewhat distinct, focusing on different consumer products and services banks provide. The five plots for each topic show the top ten words appearing in each of the five topics on the vertical axis and their intensity based on the frequency of appearance on the horizontal axis. For example, the first topic has "credit report" and "credit card" as the most frequently appearing words, while topics two and three show "credit card" and "customer services" and "credit card" and "late fee" as the most commonly appearing words, respectively. Topics four and five stand out as they correspond to loans and customer accounts, with "loan modification" and "checking account" as the most frequently appearing words. The distinct nature of these topics highlights the range of consumer complaints in the database as well as bank services that receive the most complaints.

We repeat the same exercises and ask ChatGPT to classify the topics instead, instructing it to also produce a classification with 5 main topics. Unlike traditional LDA, however, we do not apply pre-processing of data by removing bank names. Figure 12, Panel B presents topics generated by ChatGPT. There are several differences between topics generated by LDA and ChatGPT. In particular, ChatGPT-generated topics can be classified as (i) mortgage reporting, (ii) customer service, (iii) customer accounts, (iv) credit reporting, and (v) insurance and mortgage reporting. Both classifications of topics, however, confirm the prevalence of complaints relating to consumer accounts, which we find to have an important association with the decline in deposits in our earlier analyses.

[22] We construct a measure of "topic intensity" based on this analysis and are implementing these additional analyses in a regression model framework.

5. Conclusion

This study offers novel evidence on the role of consumer complaint disclosures as a mechanism for market discipline in the banking sector. Our analysis of the CFPB consumer complaint database reveals a significant relationship between the public disclosure of complaints and changes in bank deposits, especially uninsured deposits.

Specifically, we find that banks subject to CFPB oversight and receiving a higher number of consumer complaints experience a decline in uninsured deposits, suggesting that depositors actively monitor bank performance and respond to negative information. This effect is more pronounced for banks with a higher intensity of complaints. Furthermore, we find that banks increase offered deposit rates, particularly for longer-term deposits. This indicates an attempt to attract new deposits and mitigate the negative impact of declines in deposits. These findings support the hypothesis that the disclosure of consumer complaints serves as a disciplinary mechanism, influencing both depositor behavior and bank pricing strategies.

Our analysis extends beyond simply counting the number of complaints. We examined the text of the complaints, leveraging AI-powered textual analysis (using ChatGPT) to create measures of narrative sophistication and resolution expectations. This provides a more nuanced understanding of the complaints and allows us to examine the effect of the *content* of complaints, revealing that complaints relating to consumer accounts lead to larger deposit outflows. Conversely, complaints containing information on resolution expectations have a positive effect on uninsured deposits. These findings underscore the importance of both the quantity and the qualitative aspects of complaints in understanding depositor reactions and resulting market discipline.

Our findings have significant implications for prudential regulators. The CFPB's role in collecting and publicly disclosing consumer complaints provides an important channel for market discipline. However, this mechanism is not uniform in its effect and is influenced by factors such as the nature and quantity of complaints, the bank's specific circumstances, and the level of consumer trust in the banking system. In an era where data plays a crucial role, the strategic utilization of complaints data can unveil patterns and trends, offering regulators a preemptive lens through which potential systemic risks can be identified and mitigated. For instance, the spike in complaints against Wells Fargo, as shown in Figure 9, helped prudential supervisors unveil Wells Fargo's fake accounts and impose strict regulatory constraints on the firm's future ability to grow. The development of sophisticated analytical tools to parse through complaints data can provide regulators with actionable insights, guiding more informed and effective supervisory actions.

Our study opens several avenues for future research. First, the current analysis focused primarily on banks under CFPB oversight. Further research should investigate whether and how the relationship between complaint disclosures and market discipline differs for banks outside of the CFPB's jurisdiction, or across different types of financial institutions. For instance, are similar market effects observed for credit unions or non-bank financial companies? Second, the relatively short timeframe of our data, even with the inclusion of the complete CFPB dataset, may limit our ability to fully capture long-term trends in market response. Studies utilizing a longer timeframe would improve confidence in our findings. Third, we focused primarily on deposit behavior as our primary measure of market discipline. Future research could consider the effects of complaint disclosures on other aspects of bank performance, such as lending practices, profitability, and market valuation. The use of textual analysis, including AI-assisted methods, could also be expanded to other related datasets such as social media posts or news coverage. Fourth, this study highlighted the importance of considering both the quantity and the qualitative nature of consumer complaints. Future research could focus on the development and application of more refined textual analysis methods to extract and operationalize additional information from complaint narratives, moving beyond sentiment analysis toward a deeper understanding of underlying consumer concerns. This could include the development of more sophisticated algorithms to better predict potential financial distress signals from consumer complaints. Finally, this study provides evidence of the value of integrating AI techniques in research on market discipline and ultimately prudential oversight. Future work could investigate how other advanced methods, such as machine learning, may further enhance the analysis of unstructured data in finance and economics.

Overall, our findings provide robust evidence demonstrating the importance of consumer complaints in driving market discipline in the banking sector. By leveraging a comprehensive dataset and advanced analytical techniques, including AI, we find that public disclosure of consumer complaints leads to tangible market responses from both depositors and banks. This highlights the crucial role of regulatory agencies in enhancing transparency and promoting accountability within the financial industry. Our findings suggest that the strategic integration of consumer complaint data with advanced analytical tools can be a powerful instrument for reinforcing market discipline and regulatory oversight. Addressing the identified avenues for future research could significantly advance our understanding of market mechanisms and the dynamic interplay between consumers, financial institutions, and regulatory bodies.

References

Acharya, V. V., & Mora, N. (2015). A crisis of banks as liquidity providers. *Journal of Finance*, *70*, 1–43.

Agarwal, S., Lucca, D., Seru, A., & Trebbi, F. (2014). Inconsistent regulators: Evidence from banking. *The Quarterly Journal of Economics*, *129*, 889–938.

An, B., Bushman, R., Kleymenova, A., & Tomy, R. E. (2024). Bank supervision and organizational capital: The case of minority lending. *Journal of Accounting Research*, *62*, 505–549.

Anbil, S. (2018). Managing stigma during a financial crisis. *Journal of Financial Economics*, *130*, 166–181.

Begley, T. A., & Purnanandam, A. (2021). Color and credit: Race, regulation, and the quality of financial services. *Journal of Financial Economics*, *141*, 48–65.

Bybee, J. L. (2023). The ghost in the machine: Generating beliefs with large language models. Working Paper.

Chen, Q., Goldstein, I., Huang, Z., & Vashishtha, R. (2022). Bank transparency and deposit flows. *Journal of Financial Economics*, *146*, 475–501.

Diamond, D., & Dybvig, P. (1983). Bank runs, deposit insurance, and liquidity. *Journal of Political Economy*, *91*, 401–419.

Dou, Y., Hung, M., She, G., & Wang, L. L. (2024). Learning from peers: Evidence from disclosure of consumer complaints. *Journal of Accounting and Economics*, *77*, 101620.

Dou, Y., & Roh, Y. (2023). Public disclosure and consumer financial protection. *Journal of Financial and Quantitative Analysis*, (pp. 1–56).

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

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

Fuster, A., Plosser, M. C., & Vickery, J. I. (2021). Does CFPB oversight crimp credit? CEPR Discussion Paper No. DP15681.

Goldstein, I., & Sapra, H. (2014). Should banks' stress test results be disclosed? An analysis of the costs and benefits. *Foundations and Trends in Finance*, *8*, 1–54.

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

Hayes, R. M., Jiang, F., & Pan, Y. (2021). Voice of the customers: Local trust culture and consumer complaints to the CFPB. *Journal of Accounting Research*, *59*, 1077–1121.

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

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

Kim, A., Muhn, M., & Nikolaev, V. (2023). From transcripts to insights: Uncovering corporate risks using generative AI. Working Paper.

Kim, A., Muhn, M., & Nikolaev, V. (2024a). Financial statement analysis with large language models. Working Paper.

Kim, A., Muhn, M., & Nikolaev, V. V. (2024b). Bloated disclosures: can chatgpt help investors process information? *Chicago Booth Research Paper*, (pp. 2023–59).

Kleymenova, A. (2018). Consequences of mandated bank liquidity disclosures. Working paper.

Kleymenova, A., & Tomy, R. E. (2022). Observing enforcement: Evidence from banking. *Journal of Accounting Research*, *60*, 1583–1633.

Mazur, L. C. (2022). *Are the voices of customers louder when they are seen? Evidence from CFPB complaints*. Ph.D. thesis.

Passalacqua, A., Angelini, P., Lotti, F., & Soggia, G. (2019). The real effects of bank supervision: Evidence from on-site bank inspections. Working Paper.

Appendix A. Variable Definitions

Variable	**Definition**	**Source**	**Code**
Dependent Variables			
Total deposits	Natural logarithm of total bank deposits.	Call Reports	RCON2200
Insured deposits	Natural logarithm of total insured bank deposits.	Call Reports	RCON2200 - (RCON2710 + RCONF051)
Uninsured deposits	Natural logarithm of total uninsured bank deposits.	Call Reports	RCON2710 + RCONF051
CD deposit rate	Natural logarithm of deposit rate on CD products.	RateWatch	
Independent Variables			
Total complaints	Total number of complaints received by bank i in quarter t	Complaints database.	
Account complaints	Total number of bank account-related complaints filed against bank i. Bank account-related complaints defined as complaints related to "Bank account or service" or "Checking or savings account" products.	Complaints database	
Complaint narratives	Proportion of complaints with published narratives relative to total number of complaints.	Complaints database	
Account narratives	Proportion of bank account-related complaints with narratives relative to total number of complaints.	Complaints database	
Public disclosure	Indicator variable that takes a value of 1 if there is at least one publicly disclosed complaint against bank i in county j in quarter t.	Complaints database	

Variable	Definition	Source	Code
High complaint	Indicator variable that takes a value of 1 if bank i received above the median number of complaints in county j in quarter t.	Complaints database	
Total complaints	Total number of complaints filed against bank i in county j in quarter t.	Complaints database	
Sophistication score	Continuous variable generated by ChatGPT assessing the sophistication of the conusmer complaint narrative	Complaints database	ChatGPT generated metric
Resolution expectation	Continuous variable generated by ChatGPT assessing the likelihood of resolution based on the consumer complaint narrative	Complaints database	ChatGPT generated metric
Control variables			
Bank size	Natural logarithm of total bank assets	Call Reports	RCON2170
Liquidity ratio	Sum of cash and balances divided by total assets	Call Reports	(RCON0071 + RCON0081) / RCON2170
NPL ratio	Sum of non-performing loans (90 or more days delinquent and nonaccrual status loans) divided by total assets	Call Reports	(RCON1403 + RCON1407) / RCON2170
ROA	Net income divided by total assets	Call Reports	RIAD4340 / RCON2170
T1 capital ratio	Tier 1 capital divided by total assets	Call Reports	(RCOA8274 + RCFA8274 + RCON8274) / RCON2170
Population	County-level population estimate	Census Bureau	
Median household income	County-level median household income estimates	Bureau of Labor Statistics	
Unemployment rate	County-level unemployment rate	Bureau of Labor Statistics	

Figure 1: Timeline of CFPB Complaints Database Changes

This figure presents a timeline of important dates related to the CFPB complaints database. The CFPB began accepting consumer complaints about credit cards on July 21, 2011. Over the course of 2011, the CFPB began accepting complaints regarding mortgages, bank accounts and services, private student loans, and consumer loans. Through 2017, the CFPB continued to follow a phased-in approach to expand the types of complaints it accepts. As of 2023, the CFPB accepts complaints related to 11 categories ranging from credit cards, through student loans, to money transfers and virtual currencies. The public database was launched on June 19, 2012 and contained credit card complaints received on or after June 1, 2012. Over time, the CFPB released all credit card complaints dating back to December 1, 2011 as well as other complaints for other products received since March 1, 2012. On June 25, 2015, the CFPB began to publish consumer complaint narratives. Narratives are available for consumers who opt-in to disclosure.

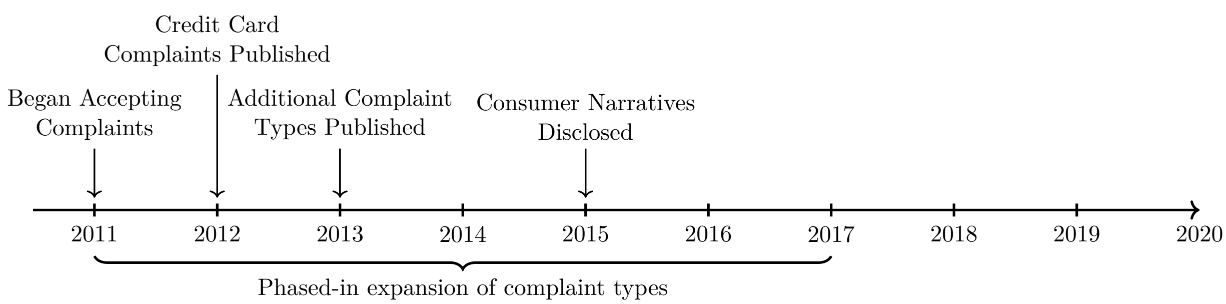


Figure 2: Sample Construction

This figure shows the distribution of CFPB supervised and non-CFPB supervised within our sample across the sample period. Our sample consists of banks with total assets reported on quarterly Call Reports between $1 and $25 billion (as of 2019 Q4). The CFPB was established in 2011 Q4. In the final quarter of our sample period, 77 banks are supervised by the CFPB and 645 are not supervised by the CFPB. Across the sample period, 79 unique banks were under CFPB supervision for at least one quarter.

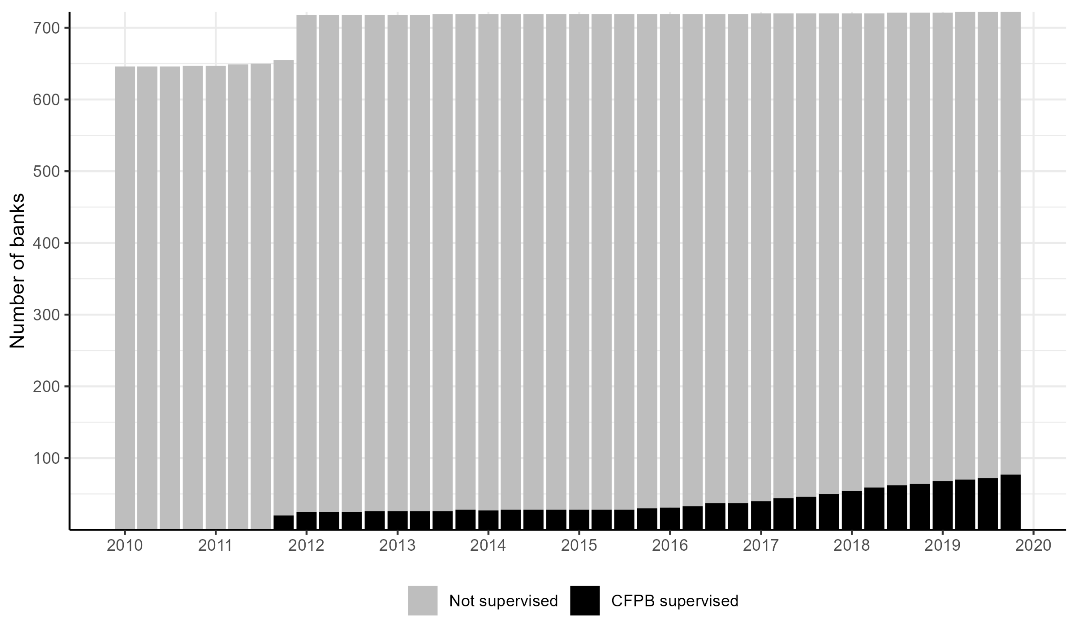


Figure 3: Geographical Distribution of Complaints

This figure shows the geographic distribution of complaints within our sample by U.S. county. Complaints tend to originate in counties near cities with larger populations (e.g., Los Angeles, Chicago, and New York). Counties shaded in grey represent counties in which there are no reported complaints or are missing county FIPS codes within our sample.

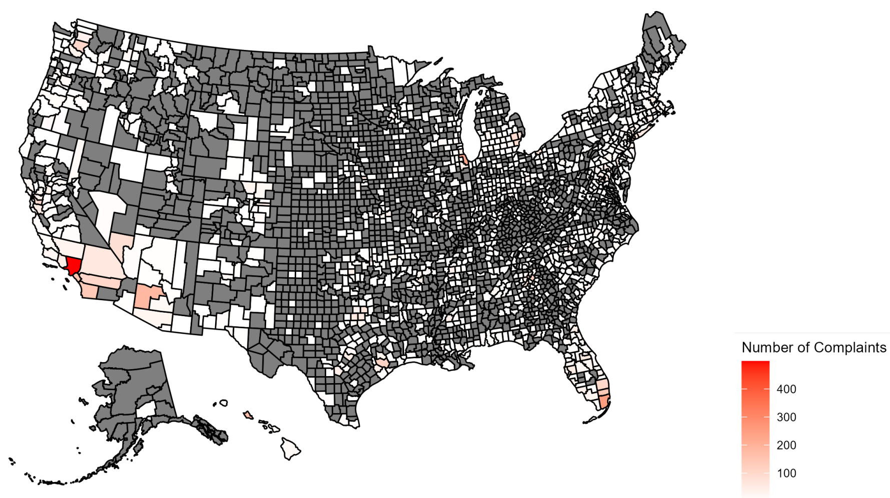
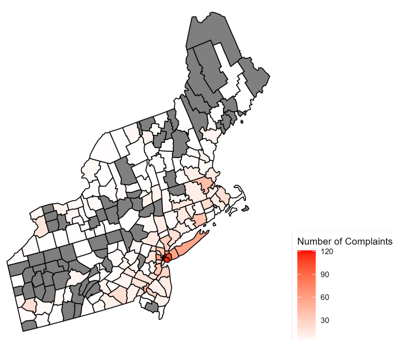


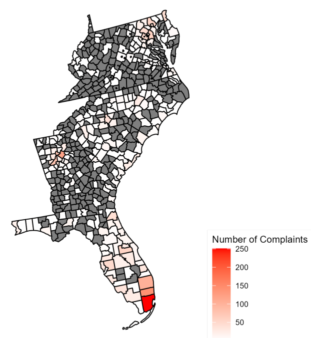
Figure 4: Regional Distribution of Complaints

This figure shows the distribution of complaints in our sample by geographical subregions. Panel A focuses on the Northeast portion of the United States and displays complaints originating from counties in PA, NY, MA, RI, CT, VT, NH, and ME. Panels B, C, and D show the same distribution for the Southeast, Midwest, and West regions, respectively. In general, complaints originate from major metropolitan areas such as New York City, Miami, Chicago, and Los Angeles. Counties shaded in grey represent counties in which there are no reported complaints or are missing county FIPS codes within our sample.

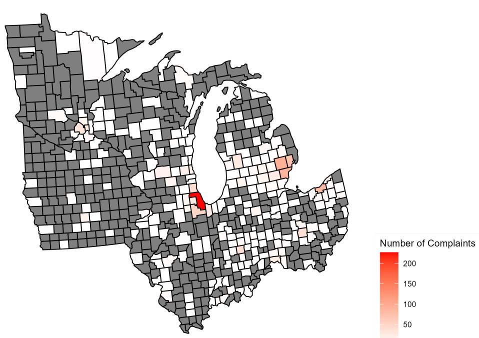
(a) Northeast



(b) Southeast



(c) Midwest



(d) West

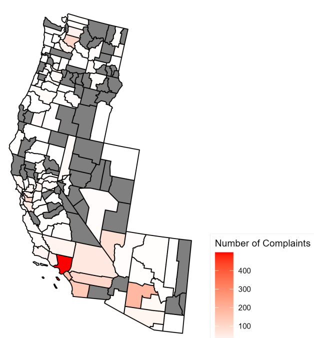


Figure 5: Consumer Complaint Trends

This figure shows the number of complaints filed against financial institutions between 2011 Q4 and 2019 Q4. Financial institutions are defined as entities that appear in FFIEC active and closed banks or the FDIC Summary of Deposits datasets. The red line is the best-fit trend line across the period. Overall, the number of complaints received by the CFPB has increased steadily since the agency's inception in 2011 Q4.

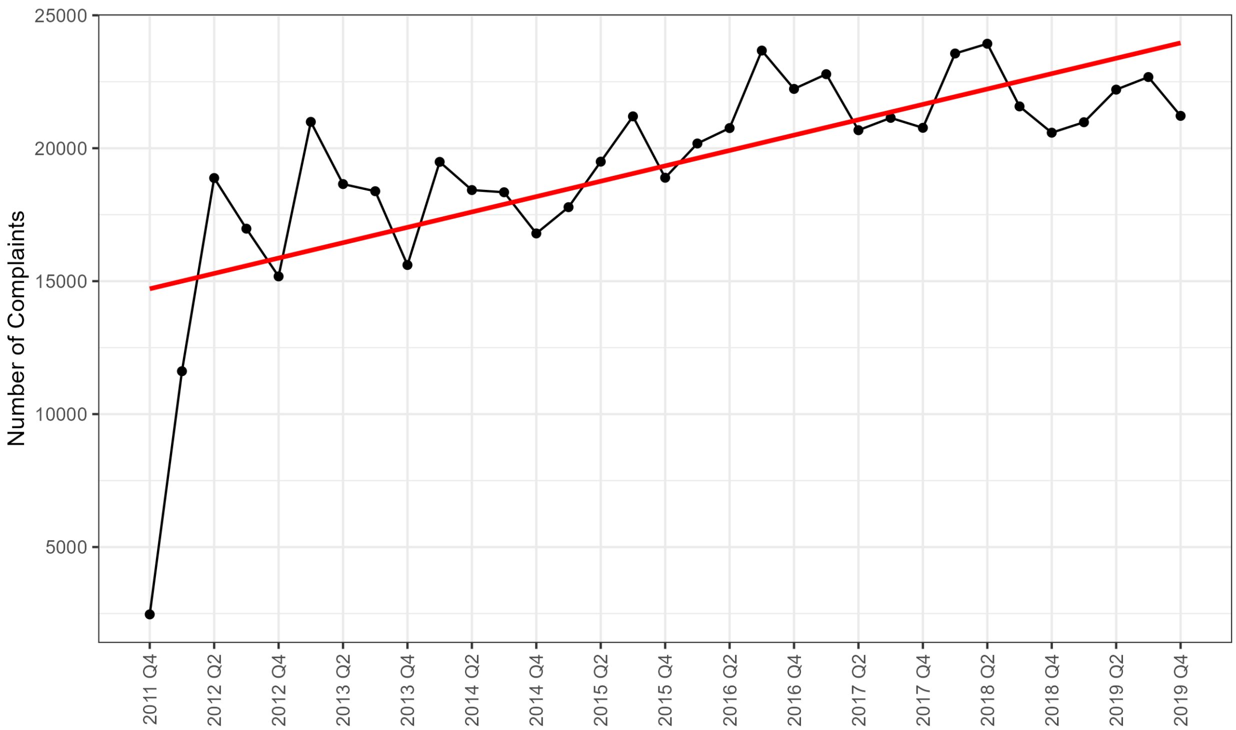


Figure 6: Complaint Trends by Product

This figure displays the time-series of total complaints filed against financial institutions between 2011 Q4 and 2019 Q4 split by product type. Financial institutions are defined as entities that appear in FFIEC active and closed banks or the FDIC Summary of Deposits datasets. The blue line plots the quarterly total of account-related complaints. Account-related complaints are defined as products in the “Bank account or service” or “Checking or savings account” categories. The red line plots the quarterly total of mortgage-related complaints. The yellow line plots the quarterly total of all other product types. Overall, the number of account-related complaints trends in parallel with other complaint categories whereas the number of mortgage-related complaints declines over time.

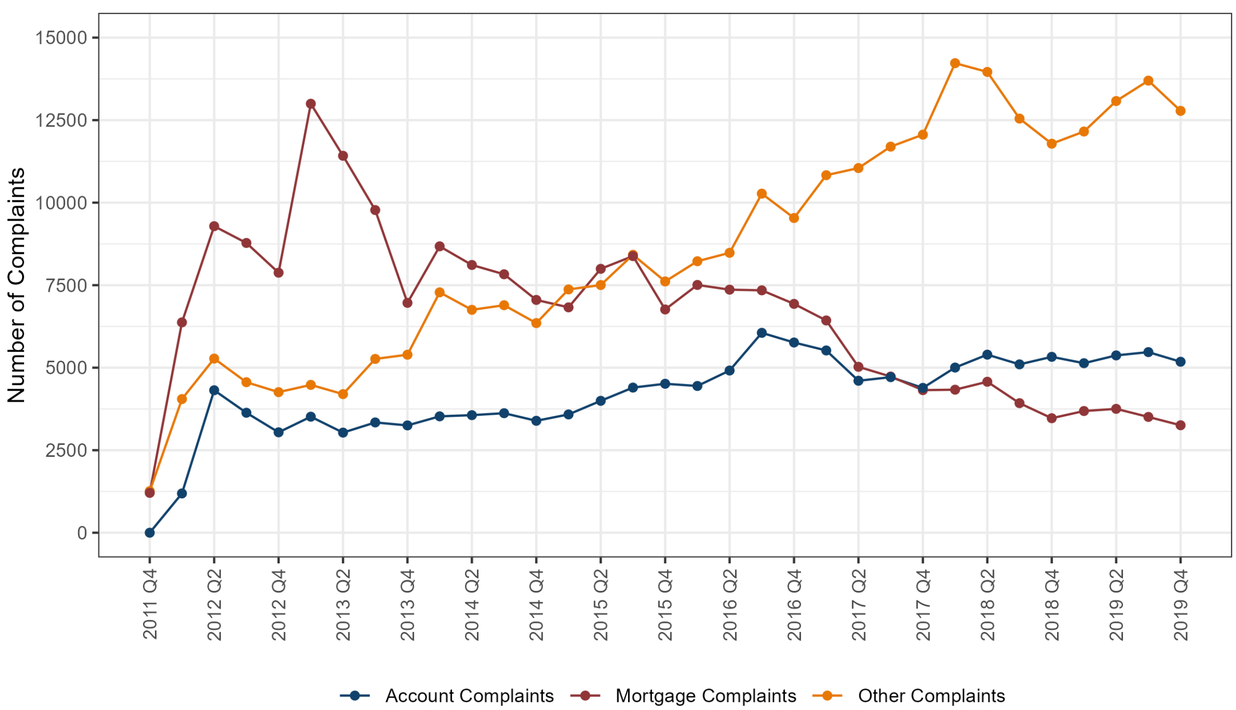


Figure 7: ChatGPT Sophistication Index by Product Type

This figure presents the average sophistication index score by product type. The sample includes 2,669 complaints with consumer narratives from 79 banks from 2015 Q1 to 2019 Q4. The sophistication index was generated by ChatGPT and assesses the complexity and quality of the narrative. The measure reflects varying levels of regulatory awareness, financial literacy, and communication skills, making them valuable for analyzing consumer behavior and institutional accountability.

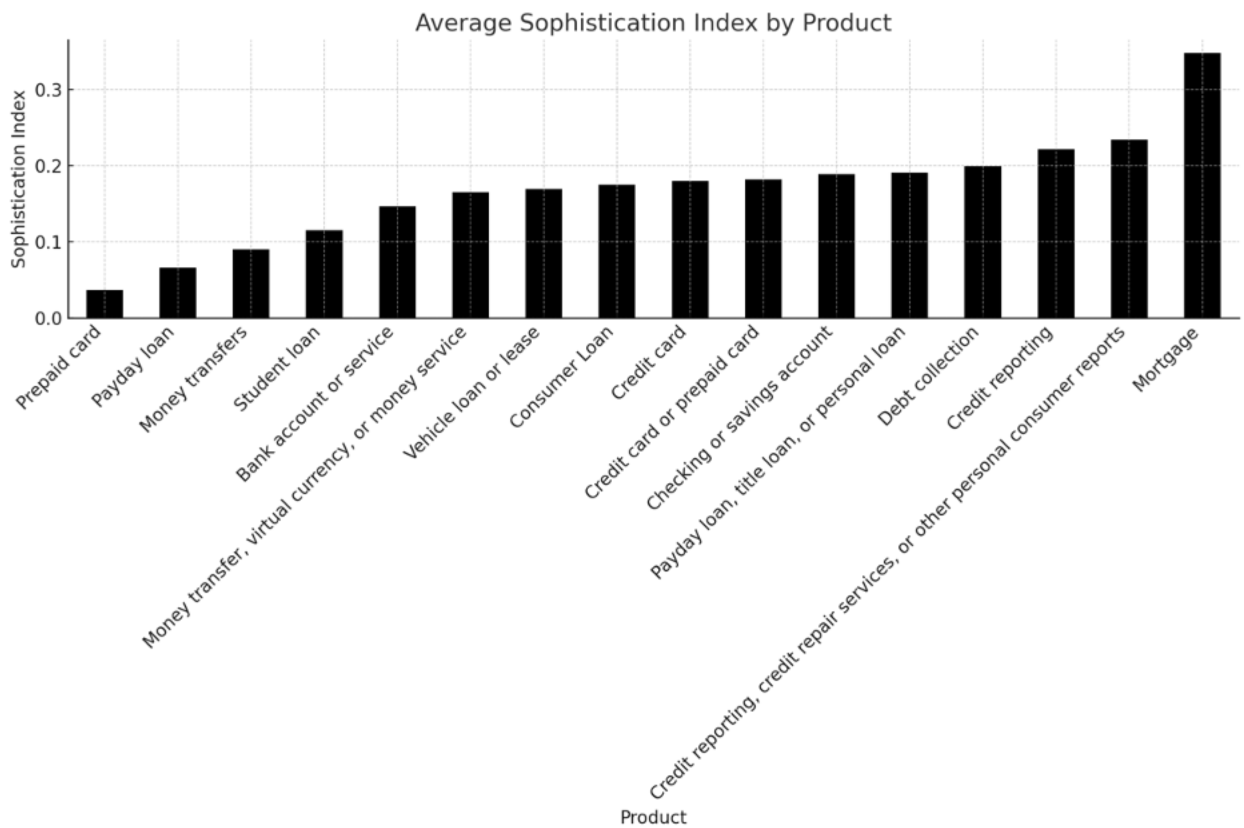


Figure 8: ChatGPT Resolution Expectation Measure by Product Type

This figure presents the average resolution expectation metric by product type. The sample includes 2,669 complaints with consumer narratives from 79 banks from 2015 Q1 to 2019 Q4. The resolution expectation metric was generated by ChatGPT and measures the clarity and specificity with which consumers articulate their desired resolutions, such as refunds or corrective actions.

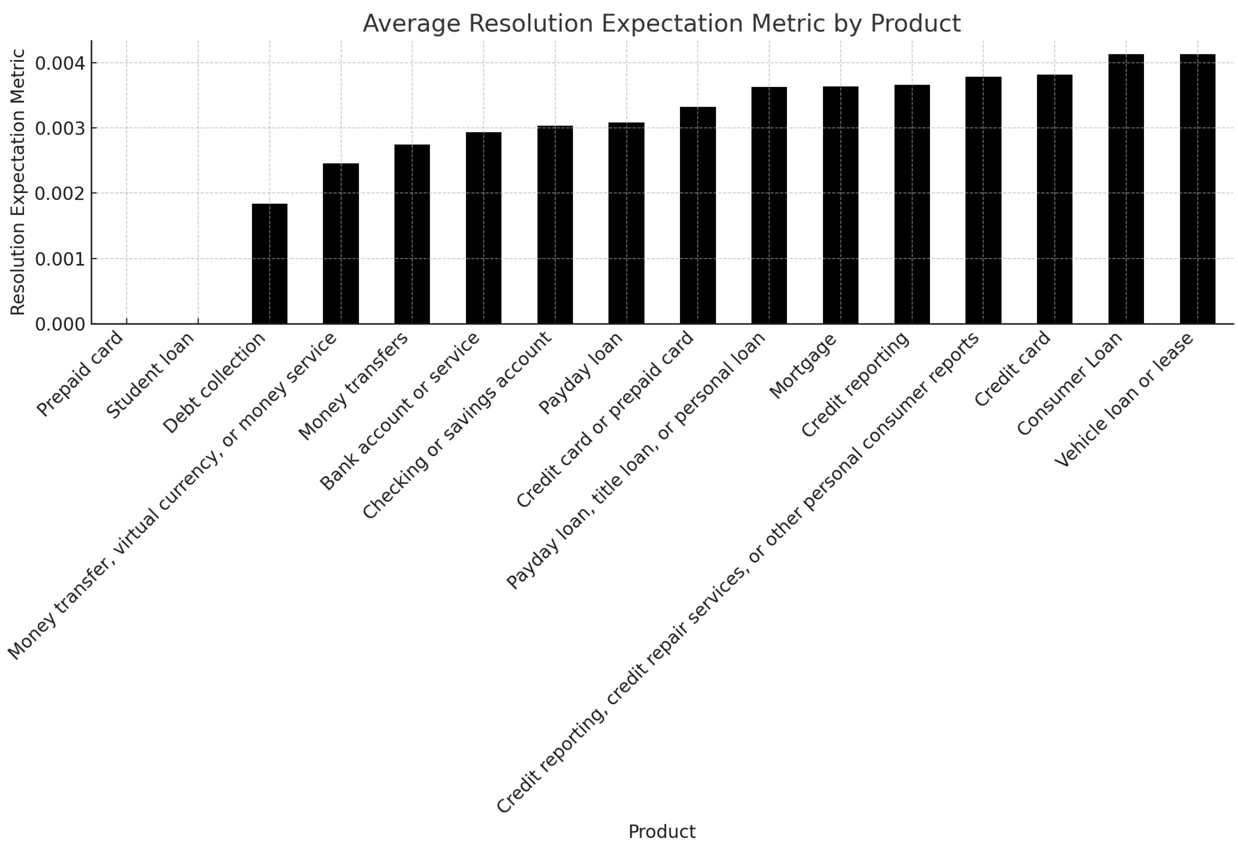


Figure 9: Bank Account-Related Complaints against Wells Fargo

This figure plots the quarterly total number of bank account-related complaints filed against Wells Fargo from 2011 Q1 to 2019 Q1. Bank account-related complaints are defined as products in the "Bank account or service" or "Checking or savings account" categories. The shaded region between 2016 Q2 and 2016 Q4 represents the time window of the Wells Fargo fake account scandal.

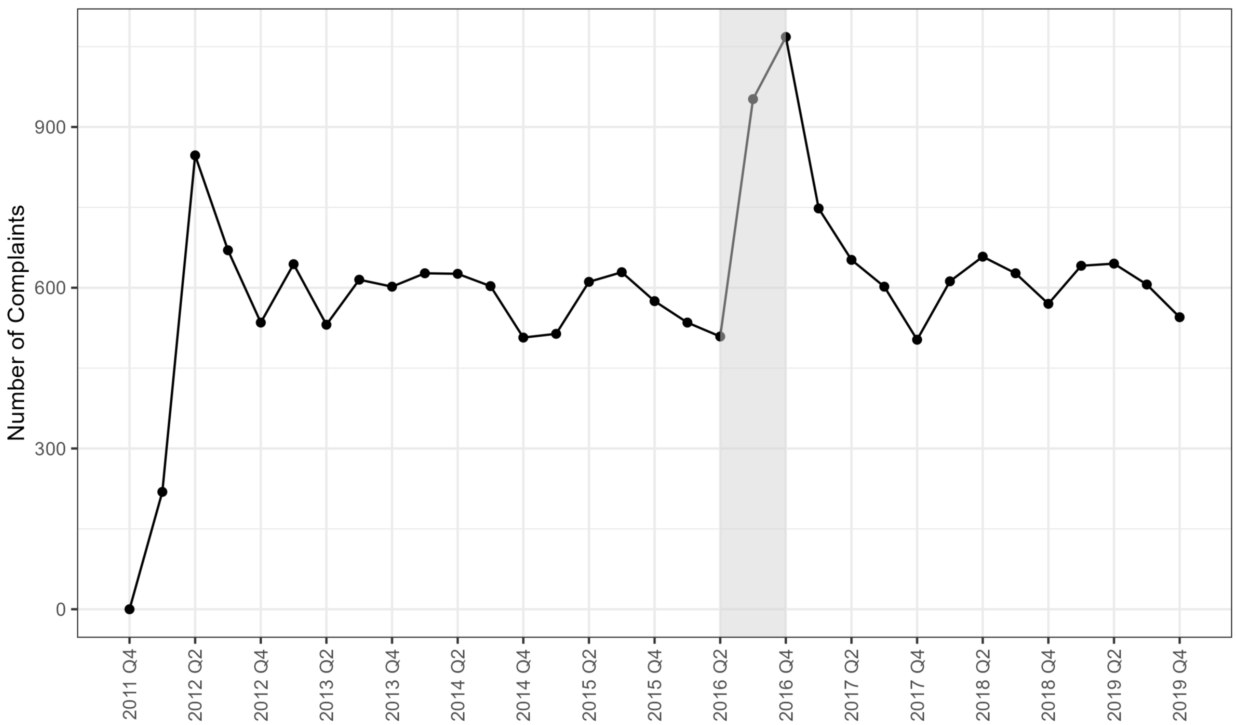


Figure 10: Credit Card-Related Complaints against Citibank

This figure plots the quarterly number of credit card-related complaints filed against Citibank. Credit card-related complaints include products classified as “Credit card” and “Credit card or prepaid card” in the CFPB complaints database. The shaded region between 2016 Q1-2016 Q3 represents the time window of Citigold reward miles issues. The shaded region between 2020 Q1-2020 Q3 represents the onset of the COVID-19 pandemic.

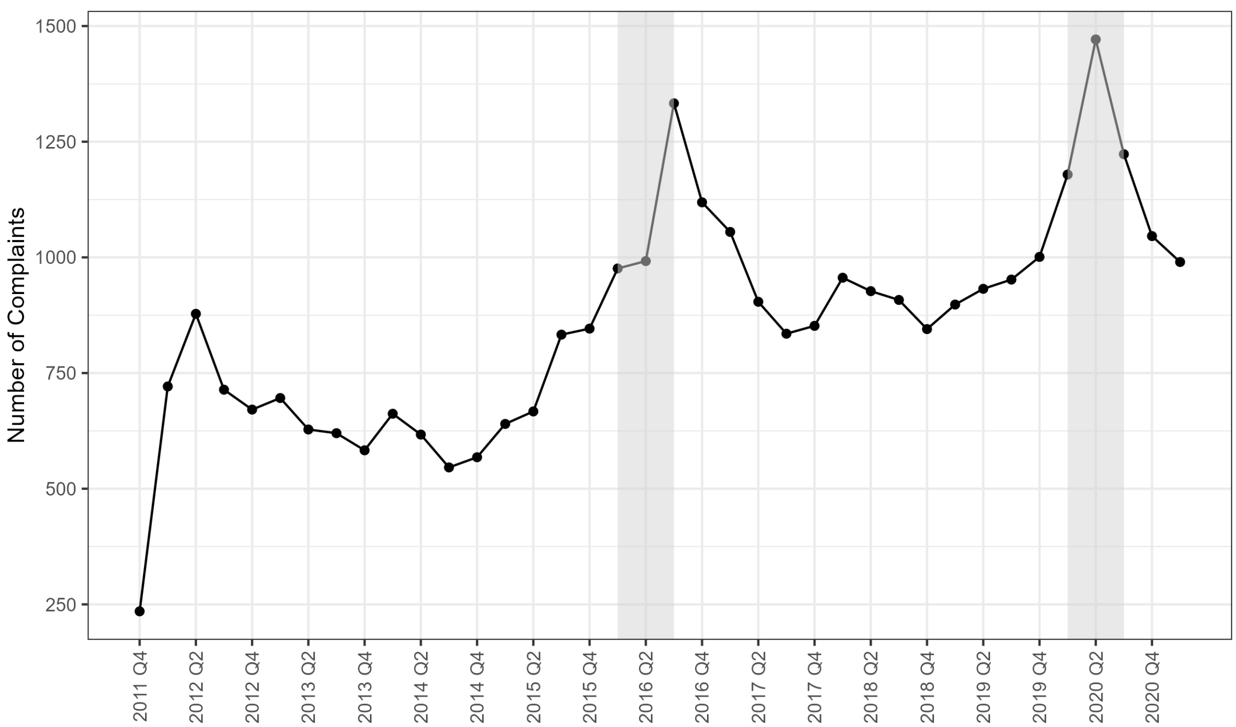


Figure 11: Credit Report-Related Complaints against Capital One

This figure plots credit report-related complaints filed against Capital One from 2012 Q1 to 2021 Q1. Credit report-related complaints are classified using the following consumer-reported product types: "Credit reporting" and "Credit reporting, credit repair services, or other personal consumer reports." The dashed line in 2017 Q1 represents the Equifax data breach. The dashed line in 2019 Q3 represents the Capital One data breach.

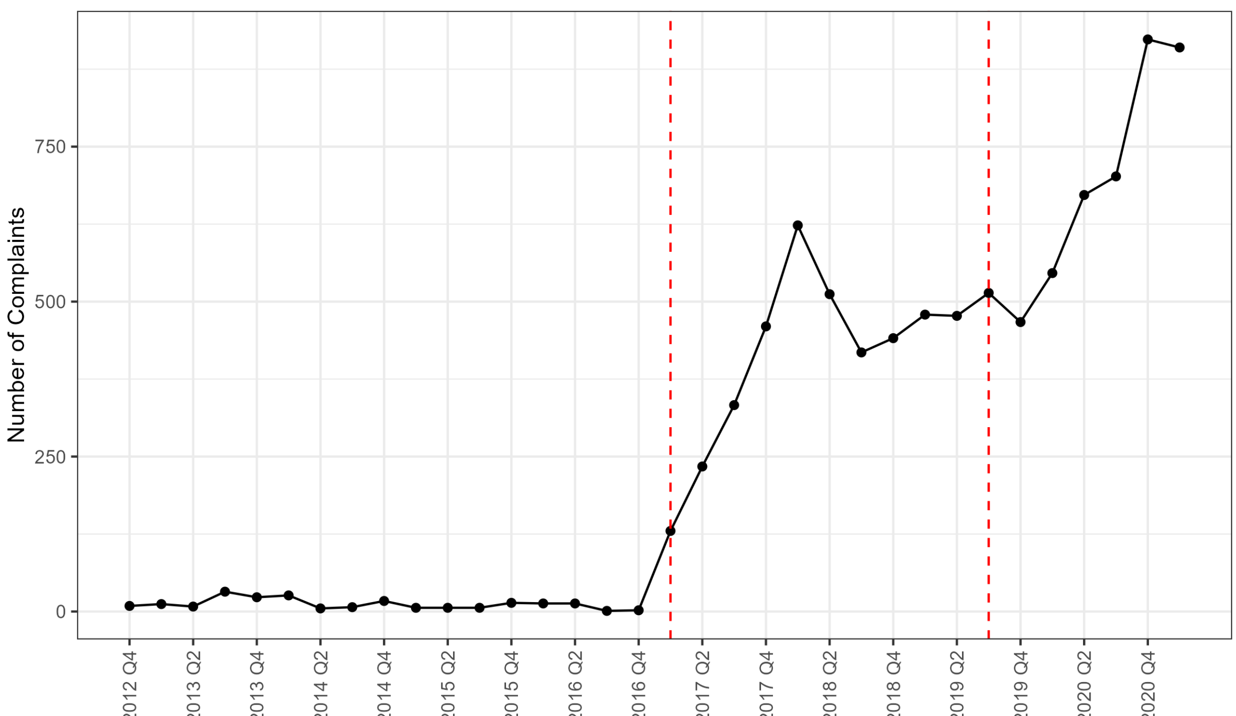
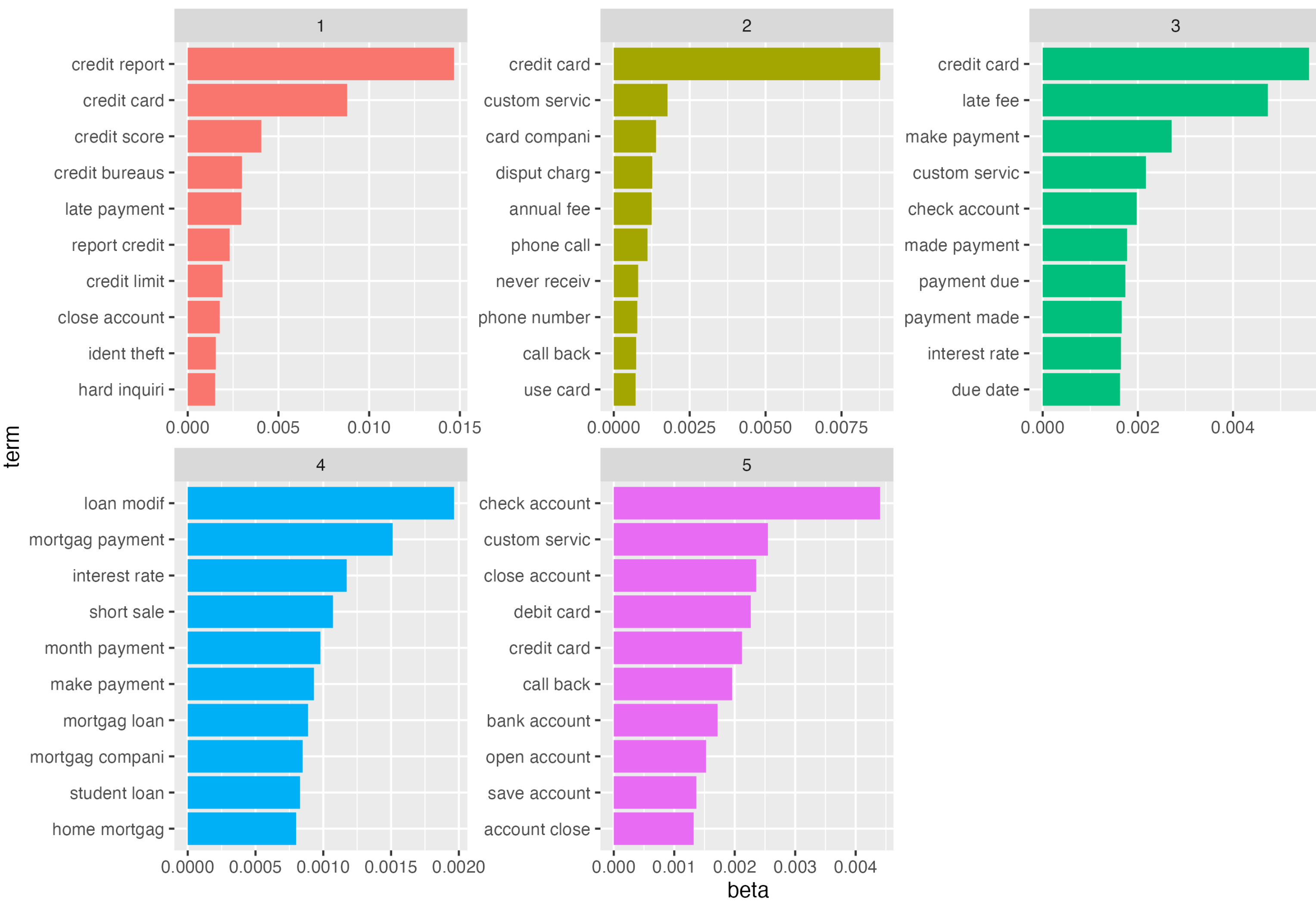


Figure 12: Comparison of Topic Analysis Results

This figure presents topic analysis results between traditional Latent Dirichlet allocation (LDA) analysis and ChatGPT generated topic analysis. Panel A presents the top 5 topics and keywords for LDA. Panel B presents the ChatGPT generated topics.

Panel A: LDA-generated Topics



Panel B: ChatGPT-generated Topics

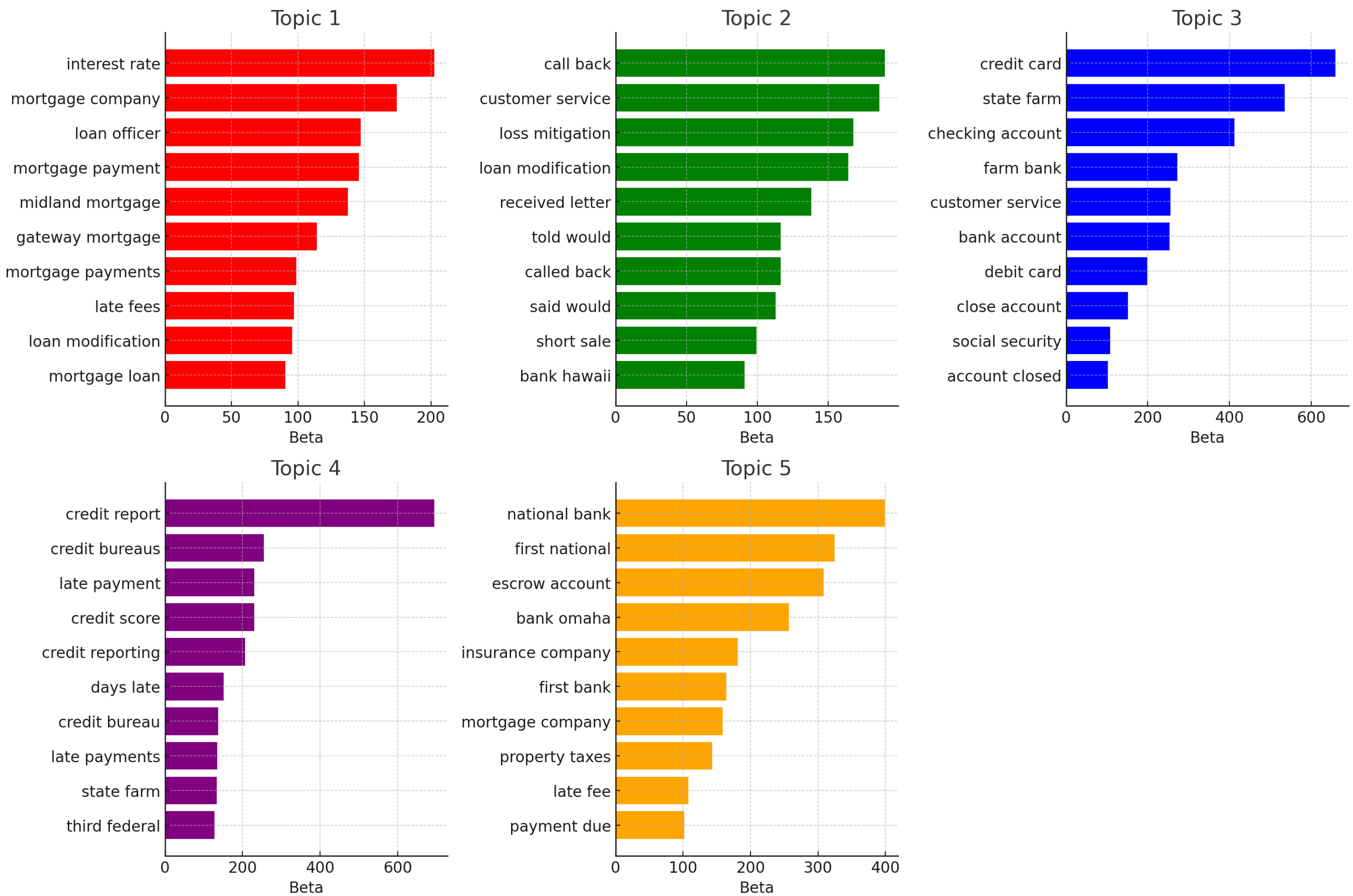


Table 1: Distribution of Complaints by Bank Assets

This table shows the distribution of complaints from 2010 Q1 to 2019 Q4. Bank assets are determined using data as of 2019 Q4. Complaints are limited to those issued against financial institutions that file quarterly Call Reports. The requirement for CFPB supervision is $10 billion in total assets. Complaints issued to banks with fewer than $10 billion assets are affiliate banks. The majority of complaints are filed against banks over $25 billion in assets. In general, larger banks receive more complaints. In total, our sample contains approximately 2% of all complaints filed against financial institutions.

	$1B-$5B	$5B-$10B	$10B-$15B	$15B-$20B	$20B-$25B	$25B+	Total
Complaints	116	22	2,530	2,432	4,952	516,685	526,737

Table 2: Descriptive Statistics

This table presents summary statistics for the 722 banks in our sample. The sample period spans 2010 Q1 to 2019 Q4. Logged variables in the regressions are presented in unlogged form. To mitigate the effect of outlier observations, variables are winsorized at the 1st and 99th percentiles. There are banks with more than $10 billion assets in the non-CFPB supervised because the sample period begins before the CFPB was established. These banks become CFPB supervised starting in 2011 Q4, after the establishment of the CFPB. All variables are defined in Appendix A.

Variable	Min.	p1	p25	Median	Mean	p75	p99	Max.	SD
CFPB supervised banks:									
Bank size ($m)	71.26	840.75	9,003.64	12,559.98	11,118.80	15,201.64	15,308.52	15,308.52	4,489.57
Deposits ($m)	17.62	17.62	6,125.46	9,645.42	8,311.34	11,638.84	11,638.84	11,638.84	3,720.80
Insured deposits ($m)	4.13	4.13	4,078.83	5,544.56	4,988.04	7,344.66	7,344.66	7,344.66	2,412.66
Liquidity ratio	0.00	0.00	0.02	0.04	0.07	0.08	0.35	0.35	0.09
NPL ratio	0.00	0.00	0.00	0.01	0.01	0.01	0.06	0.06	0.02
ROA	-0.009	-0.003	0.002	0.003	0.004	0.004	0.015	0.015	0.003
T1 capital ratio	0.06	0.06	0.09	0.10	0.12	0.13	0.32	0.32	0.05
Uninsured deposits ($m)	0.50	0.50	985.99	3,726.67	3,177.85	5,398.57	5,398.57	5,398.57	2,002.24
Non-CFPB supervised banks:									
Bank size ($m)	71.26	202.03	868.85	1,260.23	2,029.45	2,311.05	9,654.61	15,308.52	2,021.91
Deposits ($m)	17.62	126.02	718.60	1,040.19	1,628.17	1,879.29	7,339.04	11,638.84	1,555.04
Insured deposits ($m)	4.13	24.55	432.84	649.05	977.47	1,136.08	4,589.92	7,344.66	950.92
Liquidity ratio	0.00	0.01	0.02	0.04	0.06	0.08	0.35	0.35	0.06
NPL ratio	0.00	0.00	0.00	0.01	0.01	0.01	0.06	0.06	0.01
ROA	-0.009	-0.004	0.002	0.002	0.003	0.003	0.014	0.015	0.002
T1 capital ratio	0.06	0.06	0.09	0.10	0.10	0.11	0.23	0.32	0.03
Uninsured deposits ($m)	0.50	1.16	206.75	374.81	647.05	760.04	3,821.92	5,398.57	757.00

Table 3: Descriptive Statistics for ChatGPT Measures

This table presents summary statistics of the textual analysis measures generated by ChatGPT. The sample includes 79 banks that were ever supervised by the CFPB. The sample period begins in 2015 Q1, the first quarter the CFPB started collecting complaint narratives, and to 2019 Q4, the end of our sample period. In total, there are 2,669 complaints with narratives for the 79 banks in our sample across our sample period. All variables are defined in Appendix A.

Variable	Min.	p1	p25	Median	Mean	p75	p99	Max.	SD
Action Orientation	0.00	0.00	0.00	0.00	0.31	0.40	2.64	7.14	0.62
Escalation Tendency	0.00	0.00	0.00	0.00	0.03	0.00	0.68	4.00	0.18
Frustration Urgency	0.00	0.00	0.00	0.00	0.05	0.00	0.90	4.76	0.25
Resolution Expectation	0.00	0.00	0.00	0.00	0.32	0.43	2.91	7.14	0.65
Sophistication Index	0.00	0.02	0.10	0.14	0.24	0.36	0.81	0.90	0.18

Table 4: Effect of Consumer Complaints on Deposits

This table presents OLS estimates of the effect of total consumer complaints on deposits. The sample period is from 2010 Q1 to 2019 Q4. The sample includes 722 banks with assets between $1 and $25 billion in total assets as of 2019 Q4. *Total complaints* is the total number of complaints received by bank b in quarter t. Robust standard errors clustered by bank and quarter. All variables are defined in Appendix A. ***, **, and * represent statistical significance at the 1%, 5%, and 10% levels, respectively.

	Total Deposits	Insured Deposits	Uninsured Deposits
	(1)	(2)	(3)
Total complaints	-0.001	0.001	-0.002*
	(0.001)	(0.001)	(0.001)
Bank size	0.922***	0.845***	0.965***
	(0.028)	(0.052)	(0.044)
Liquidity ratio	0.083	0.082	0.010
	(0.271)	(0.623)	(0.351)
NPL ratio	-0.335	-0.931	-3.447***
	(0.554)	(2.317)	(0.911)
ROA	-2.235	-7.386	5.130
	(4.829)	(6.781)	(4.927)
T1 capital ratio	-1.976**	-2.955**	-1.215
	(0.869)	(1.312)	(1.190)
Bank FE	Yes	Yes	Yes
Quarter FE	Yes	Yes	Yes
Observations	28,136	28,136	28,136
Adjusted R^2	0.964	0.911	0.946

Table 5: Effect of Account-Related Complaints on Deposits

This table presents OLS estimates of the effect of bank account-related complaints on deposits. The sample period is from 2010 Q1 to 2019 Q4. The sample includes 722 banks with assets between \$1 and \$25 billion in total assets as of 2019 Q4. Bank account-related complaints are classified as complaints in the "Bank account or service" and "Checking or savings account" product categories. *Account complaints* is the proportion of bank account-related complaints relative to all complaints. Bank controls are lagged by one quarter. Robust standard errors clustered by bank and quarter. All variables are defined in Appendix A. ***, **, and * represent statistical significance at the 1%, 5%, and 10% levels, respectively.

	Total Deposits	Insured Deposits	Uninsured Deposits
	(1)	(2)	(3)
Account complaints	-0.042**	-0.011	-0.066**
	(0.018)	(0.024)	(0.032)
Bank size	0.922***	0.845***	0.966***
	(0.028)	(0.052)	(0.044)
Liquidity ratio	0.081	0.082	0.008
	(0.271)	(0.623)	(0.350)
NPL ratio	-0.351	-0.923	-3.472***
	(0.556)	(2.320)	(0.912)
ROA	-2.207	-7.402	5.174
	(4.823)	(6.779)	(4.921)
T1 capital ratio	-1.978**	-2.952**	-1.218
	(0.868)	(1.312)	(1.188)
Bank FE	Yes	Yes	Yes
Quarter FE	Yes	Yes	Yes
Observations	28,136	28,136	28,136
Adjusted R^2	0.964	0.911	0.946

Table 6: Effect of Total Complaints on Deposits by Bank Profitability

This table presents OLS estimates of the effect of total consumer complaints on deposits, partitioned by high and low ROA. The sample period is from 2010 Q1 to 2019 Q4. The sample includes 722 banks with assets between \$1 and \$25 billion in total assets as of 2019 Q4. *Total complaints* is the total number of complaints received by bank b in quarter t. The p-values of the Wald statistic testing for differences in coefficients between columns (1)-(4), (2)-(5), and (3)-(6) are 0.223, 0.416, and 0.371, respectively. Bank controls are lagged by one quarter. Robust standard errors clustered by bank and quarter. All variables are defined in Appendix A. ***, **, and * represent statistical significance at the 1%, 5%, and 10% levels, respectively.

	High ROA			Low ROA		
	Total Deposits	Insured Deposits	Uninsured Deposits	Total Deposits	Insured Deposits	Uninsured Deposits
	(1)	(2)	(3)	(4)	(5)	(6)
Total complaints	-0.002	0.000	-0.002	-0.001	0.002	-0.003*
	(0.001)	(0.002)	(0.001)	(0.001)	(0.001)	(0.001)
Bank size	0.879***	0.806***	0.921***	0.946***	0.867***	0.985***
	(0.044)	(0.077)	(0.054)	(0.016)	(0.045)	(0.045)
Liquidity ratio	0.395	0.933	0.349	-0.136	-0.555*	-0.174
	(0.357)	(1.065)	(0.397)	(0.139)	(0.323)	(0.303)
NPL ratio	-1.181	-4.691	-3.574***	0.007	0.896	-2.597**
	(0.820)	(4.982)	(1.196)	(0.373)	(0.728)	(1.094)
ROA	-1.572	-6.275	3.205	-3.207	-8.469	3.289
	(8.079)	(11.330)	(7.809)	(3.284)	(5.079)	(3.745)
T1 capital ratio	-2.104*	-2.875*	-1.167	-1.336***	-2.325	-0.529
	(1.182)	(1.481)	(1.401)	(0.472)	(1.557)	(0.750)
Bank FE	Yes	Yes	Yes	Yes	Yes	Yes
Quarter FE	Yes	Yes	Yes	Yes	Yes	Yes
Observations	14,070	14,070	14,070	14,066	14,066	14,066
Adjusted R^2	0.958	0.897	0.948	0.979	0.941	0.951

Table 7: Effect of Account-Related Complaints on Deposits by Bank Profitability

This table presents OLS estimates of the effect of bank account-related complaints on deposits, partitioned by high and low ROA. The sample period is from 2010 Q1 to 2019 Q4. The sample includes 722 banks with assets between $1 and $25 billion in total assets as of 2019 Q4. Bank account-related complaints are classified as complaints in the "Bank account or service" and "Checking or savings account" product categories. *Account complaints* is the proportion of bank account-related complaints relative to all complaints. The *p*-values of the Wald statistic testing for differences in coefficients between columns (1)-(4), (2)-(5), and (3)-(6) are 0.823, 0.950, and 0.777, respectively. Bank controls are lagged by one quarter. Robust standard errors clustered by bank and quarter. All variables are defined in Appendix A. ***, **, and * represent statistical significance at the 1%, 5%, and 10% levels, respectively.

	High ROA			Low ROA		
	Total Deposits	Insured Deposits	Uninsured Deposits	Total Deposits	Insured Deposits	Uninsured Deposits
	(1)	(2)	(3)	(4)	(5)	(6)
Account complaints	-0.036	-0.022	-0.048	-0.042**	-0.025	-0.033
	(0.023)	(0.034)	(0.041)	(0.019)	(0.031)	(0.036)
Bank size	0.879***	0.806***	0.922***	0.946***	0.867***	0.986***
	(0.044)	(0.077)	(0.055)	(0.016)	(0.045)	(0.045)
Liquidity ratio	0.392	0.934	0.347	-0.138	-0.556*	-0.175
	(0.357)	(1.064)	(0.396)	(0.139)	(0.323)	(0.304)
NPL ratio	-1.203	-4.695	-3.598***	0.000	0.914	-2.624**
	(0.824)	(4.990)	(1.198)	(0.375)	(0.725)	(1.096)
ROA	-1.498	-6.284	3.271	-3.190	-8.463	3.307
	(8.062)	(11.306)	(7.795)	(3.286)	(5.084)	(3.758)
T1 capital ratio	-2.107*	-2.875*	-1.169	-1.338***	-2.322	-0.534
	(1.181)	(1.480)	(1.400)	(0.471)	(1.557)	(0.750)
Bank FE	Yes	Yes	Yes	Yes	Yes	Yes
Quarter FE	Yes	Yes	Yes	Yes	Yes	Yes
Observations	14,070	14,070	14,070	14,066	14,066	14,066
Adjusted R^2	0.958	0.897	0.948	0.979	0.941	0.951

Table 8: Effect of Complaint Narratives on Deposits

This table presents OLS estimates of the effect of complaint narratives on deposits. The sample period is from 2010 Q1 to 2019 Q4. The sample includes 722 banks with assets between $1 and $25 billion in total assets as of 2019 Q4. *Total complaints* is the total number of complaints for bank b in time t. *Complaint narratives* is the proportion of complaints with narratives relative to all complaints. Bank controls are lagged by one quarter. Robust standard errors clustered by bank and quarter. All variables are defined in Appendix A. ***, **, and * represent statistical significance at the 1%, 5%, and 10% levels, respectively.

	Total Deposits	Insured Deposits	Uninsured Deposits	Total Deposits	Insured Deposits	Uninsured Deposits
	(1)	(2)	(3)	(4)	(5)	(6)
Complaint narratives	-0.074***	-0.025	-0.088*	-0.067***	-0.040	-0.075*
	(0.024)	(0.037)	(0.046)	(0.024)	(0.036)	(0.044)
Total complaints				-0.001	0.002	-0.001*
				(0.001)	(0.001)	(0.001)
Bank size	0.922***	0.845***	0.965***	0.922***	0.845***	0.965***
	(0.028)	(0.052)	(0.044)	(0.028)	(0.052)	(0.044)
Liquidity ratio	0.082	0.082	0.009	0.082	0.081	0.010
	(0.270)	(0.623)	(0.350)	(0.271)	(0.623)	(0.351)
NPL ratio	-0.358	-0.926	-3.478***	-0.351	-0.941	-3.465***
	(0.556)	(2.322)	(0.913)	(0.555)	(2.320)	(0.912)
ROA	-2.213	-7.403	5.163	-2.224	-7.379	5.142
	(4.823)	(6.779)	(4.921)	(4.826)	(6.779)	(4.926)
T1 capital ratio	-1.981**	-2.954**	-1.222	-1.980**	-2.957**	-1.219
	(0.868)	(1.312)	(1.188)	(0.869)	(1.313)	(1.190)
Bank FE	Yes	Yes	Yes	Yes	Yes	Yes
Quarter FE	Yes	Yes	Yes	Yes	Yes	Yes
Observations	28,136	28,136	28,136	28,136	28,136	28,136
Adjusted R^2	0.964	0.911	0.946	0.964	0.911	0.946

Table 9: Effect of Account-Related Complaint Narratives on Deposits

This table presents OLS estimates of the effect of account-related complaint narratives on deposits. The sample period is from 2010 Q1 to 2019 Q4. The sample includes 722 banks with assets between \$1 and \$25 billion in total assets as of 2019 Q4. Bank account-related complaints are classified as complaints related to products under "Bank account or service" and "Checking or savings account." *Account complaints* is the proportion of bank account-related complaints relative to all complaints. *Account narratives* is the proportion of account-related complaints with narratives relative to all account-related complaints. Bank controls are lagged by one quarter. Robust standard errors clustered by bank and quarter. All variables are defined in Appendix A. ***, **, and * represent statistical significance at the 1%, 5%, and 10% levels, respectively.

	Total Deposits	Insured Deposits	Uninsured Deposits	Total Deposits	Insured Deposits	Uninsured Deposits
	(1)	(2)	(3)	(4)	(5)	(6)
Account narratives	-0.056**	-0.002	-0.080**	-0.034*	0.006	-0.044
	(0.023)	(0.030)	(0.038)	(0.019)	(0.024)	(0.034)
Account complaints				-0.031*	-0.013	-0.053
				(0.017)	(0.022)	(0.032)
Bank size	0.922***	0.845***	0.965***	0.922***	0.845***	0.965***
	(0.028)	(0.052)	(0.044)	(0.028)	(0.052)	(0.044)
Liquidity ratio	0.081	0.082	0.008	0.081	0.082	0.007
	(0.271)	(0.623)	(0.350)	(0.271)	(0.623)	(0.350)
NPL ratio	-0.351	-0.922	-3.471***	-0.353	-0.923	-3.475***
	(0.556)	(2.320)	(0.912)	(0.555)	(2.320)	(0.912)
ROA	-2.208	-7.404	5.172	-2.204	-7.402	5.178
	(4.822)	(6.779)	(4.919)	(4.822)	(6.779)	(4.920)
T1 capital ratio	-1.981**	-2.953**	-1.223	-1.980**	-2.952**	-1.221
	(0.868)	(1.313)	(1.188)	(0.868)	(1.312)	(1.188)
Bank FE	Yes	Yes	Yes	Yes	Yes	Yes
Quarter FE	Yes	Yes	Yes	Yes	Yes	Yes
Observations	28,136	28,136	28,136	28,136	28,136	28,136
Adjusted R^2	0.964	0.911	0.946	0.964	0.911	0.946

Table 10: Effect of Complaint Sophistication on Deposits

This table presents the regression results of ChatGPT-generated textual analysis of consumer complaint narratives assessing the sophistication of the narrative. *Sophistication score* is the average sophistication score of complaint narratives for bank b in quarter t. Bank controls are lagged by one quarter and include liquidity ratio, capital ratio, ROA, and log(assets) as reported on quarterly Call Reports. The sample includes 79 CFPB-supervised banks from 2015 Q1 to 2019 Q4. Bank controls are lagged by one quarter. Robust standard errors clustered by bank and quarter. All variables are defined in Appendix A. ***, **, and * represent statistical significance at the 1%, 5%, and 10% levels, respectively.

	Total Deposits	Insured Deposits	Uninsured Deposits
	(1)	(2)	(3)
Sophistication score	0.017	0.139	0.090
	(0.130)	(0.144)	(0.192)
Bank size	1.025***	1.154***	0.899***
	(0.177)	(0.187)	(0.283)
Liquidity ratio	-1.963*	-2.374**	-2.576*
	(0.940)	(1.093)	(1.425)
NPL ratio	-1.896	-4.955	7.716
	(2.493)	(3.122)	(5.327)
ROA	-23.231	-36.757	32.004
	(30.722)	(41.126)	(34.666)
T1 capital ratio	-4.505	-5.384	-7.077*
	(2.607)	(3.154)	(3.998)
Bank FE	Yes	Yes	Yes
Quarter FE	Yes	Yes	Yes
Observations	1,547	1,547	1,547
Adjusted R^2	0.909	0.880	0.907

Table 11: Effect of Complaint Resolution Expectation on Deposits

This table presents the regression results of ChatGPT-generated textual analysis of consumer complaint narratives assessing the narratives' resolution expectation. *Resolution expectation* is the average resolution expectation of complaint narratives for bank b in quarter t. Bank controls are lagged by one quarter and include liquidity ratio, capital ratio, ROA, and log(assets) as reported on quarterly Call Reports. The sample includes 79 CFPB-supervised banks from 2015 Q1 to 2019 Q4. Bank controls are lagged by one quarter. Robust standard errors clustered by bank and quarter. All variables are defined in Appendix A. ***, **, and * represent statistical significance at the 1%, 5%, and 10% levels, respectively.

	Total Deposits	Insured Deposits	Uninsured Deposits
	(1)	(2)	(3)
Resolution expectation	0.018	0.014	0.025*
	(0.013)	(0.020)	(0.014)
Bank size	1.025***	1.156***	0.900***
	(0.176)	(0.187)	(0.282)
Liquidity ratio	-1.961*	-2.364**	-2.569*
	(0.938)	(1.094)	(1.424)
NPL ratio	-1.874	-4.897	7.770
	(2.459)	(3.081)	(5.320)
ROA	-23.280	-36.599	32.036
	(30.680)	(41.061)	(34.546)
T1 capital ratio	-4.505	-5.382	-7.075*
	(2.607)	(3.153)	(3.997)
Bank FE	Yes	Yes	Yes
Quarter FE	Yes	Yes	Yes
Observations	1,547	1,547	1,547
Adjusted R^2	0.909	0.880	0.907

Table 12: Effect of Complaints on Deposit Rates

This table presents the regression results estimating log changes in bank deposit rates. Bank controls are lagged by one quarter and include liquidity ratio, capital ratio, ROA, and log(assets) as reported on quarterly Call Reports. County controls include population, median household income, and the unemployment rate, as reported by the Census Bureau and BLS. Data aggregated to bank-county (complaint)-quarter level. Bank controls are lagged by one quarter. Robust standard errors clustered by bank and quarter. All variables are defined in Appendix A. ***, **, and * represent statistical significance at the 1%, 5%, and 10% levels, respectively.

	CD Deposit Rate		
	(1)	(2)	(3)
Public disclosure	0.007***		
	(0.003)		
High Complaint		0.007***	
		(0.003)	
Total Complaints			0.003***
			(0.001)
Bank and County Controls	Yes	Yes	Yes
Quarter FE	Yes	Yes	Yes
County FE	Yes	Yes	Yes
Bank FE	Yes	Yes	Yes
Observations	153,886	153,886	153,886
Adjusted R^2	0.801	0.801	0.801

Appendix OA. Online Appendix

Appendix OA.1. Description of ChatGPT generated output

This section provides an overview of how ChatGPT was used to generate measures from consumer complaint narratives, beginning with the foundational documents and the underlying data that informed the analysis. It then details the iterative process we used to develop metrics. The section concludes with a discussion of the validation steps, and the practical applications of these measures in understanding consumer complaints.

Initialization

Two CFPB documents that are publicly available informed our approach:

1. **CFPB Narrative Scrubbing Standards**: These standards are designed to protect consumer privacy by systematically redacting sensitive information, such as names, account numbers, and other personally identifiable details, replacing them with placeholders like "XXXX." While critical for ensuring compliance with privacy regulations, this process can obscure key contextual details, such as specific transaction amounts or timelines, which may pose analytical challenges when interpreting the narratives for both LLM, and researchers.

2. **CFPB Supervision and Examination Manual**: This manual establishes comprehensive institutional guidelines for ensuring consumer protection and effective complaint resolution. It provides a framework for evaluating compliance with regulatory standards, detailing expectations for addressing consumer grievances, safeguarding financial rights, and maintaining transparency. It was used in informing dimensions such as regulatory awareness and financial terminology.

We began by providing the two documents to the LLM and inform insights in the development of key dimensions. Additionally, we analyzed a detailed Excel file containing consumer complaint narratives uploaded during the development process. This dataset allowed us to observe real-world examples of narratives, understand the variability in structure and content, and test preliminary metrics iteratively.

<u>Collaborative Process and AI Contributions</u>

The sophistication measure evolved through an iterative collaboration between us, as researchers, and a conversational AI, which provided structured methodologies, computational insights, and actionable tools. This process was grounded in:

1. **Exploration of Narrative Data**: The uploaded Excel file provided a representative sample of consumer narratives, illustrating the presence of regulatory terms, financial jargon, and redactions. By examining this dataset, we identified variability in narrative quality, which informed the dimensions of sophistication.

2. **Defining Key Dimensions**: Based on the dataset and prompts, the AI proposed dimensions such as regulatory awareness, financial terminology, and specificity. For

instance, narratives referencing laws (e.g., “Fair Credit Reporting Act”) or containing terms like “interest rate” demonstrated higher sophistication. The design and refinement of prompts were instrumental in eliciting targeted outputs from the LLM. Prompts were specific and developed to align with theoretical frameworks, ensuring that the proposed dimensions captured meaningful aspects of the narratives while avoiding overfitting.

3. **Incorporating TF-IDF Analysis**: TF-IDF (Term Frequency-Inverse Document Frequency) analysis was suggested to enhance our ability to identify key terms that characterize sophisticated narratives. This approach quantitatively determines the importance of terms in a narrative relative to the entire dataset, helping distinguish regulatory or financial terms from generic language. For example, terms like “interest rate,” “escrow account,” or “loan modification” consistently emerged as indicators of financial literacy, aligning with dimensions like regulatory awareness and financial terminology.

4. **Iterative Refinement Through Feedback**: We tested metrics on subsets of the narratives and provided feedback to refine scoring logic. For example, when clarity and specificity metrics underperformed on overly concise narratives, adjustments were made to consider tone and logical flow. For example, initial iterations revealed inconsistencies in detecting specificity for narratives with minimal content. These cases were addressed by recalibrating the model’s scoring logic and refining the prompts.

5. **Treatment of Redactions**: A significant focus of our discussions was the role of redacted text (e.g., “XXXX”). Early on, we debated whether redactions should be penalized. The related prompt is as follows:

   > *“Why should we penalize for ‘XXXX’? Did you mean high counts of ‘XXXX’ in one narrative?”*

   The AI suggested treating redactions as contextual factors rather than uniform detractors. This evolved into a nuanced discussion:

   > *“On the other hand, do redacted complaints also mean they include important information? Your thoughts?”*

   These discussions highlighted the dual role of redactions: while they obscure details, they may also signal high-stakes complaints. To account for this, the AI proposed a holistic approach, weighing redactions against other dimensions like tone, regulatory awareness, and vocabulary complexity. This approach was further refined after clarifying variations in redacted text:

   > *“Redacted can have multiple X—not just 4 Xs—did you get it?”*

Dimensions of Sophistication

The LLM suggested a final sophistication measure, which integrates the following seven dimensions. These dimensions were informed by both exploratory analysis of the dataset and iterative refinements:

1. **Regulatory Awareness**: Captures references to laws and regulatory agencies, reflecting the consumer's understanding of their rights.

2. **Financial Terminology**: Measures the use of domain-specific terms, such as "adjustable rates" or "escrow accounts," signaling financial literacy.

3. **Specificity**: Assesses the inclusion of concrete details (e.g., dollar amounts, transaction dates, and product types), adjusted for redactions.

4. **Clarity and Structure**: Evaluates readability, coherence, and logical sequencing.

5. **Professionalism and Politeness**: Considers tone and phrasing, focusing on respectful and formal communication.

6. **Vocabulary Complexity**: Analyzes linguistic richness and diversity using lexical metrics.

7. **Impact of Redactions**: Quantifies redacted text using regex patterns and contextualizes its effect on narrative clarity and informativeness.

<u>Integration of NLP, TF-IDF, and LLM Techniques</u>

The sophistication measure combines rule-based NLP methods, TF-IDF analysis, and LLM-based contextual evaluations to comprehensively assess narratives:

1. **NLP Feature Extraction**: Python and R scripts were developed to extract measurable features such as keyword counts, regex-matched details, readability scores, and redactions. For example, regex patterns were used to detect variations of redacted text (e.g., "XXXX" vs. "XXXXX"), ensuring consistent treatment.

2. **TF-IDF Analysis**: TF-IDF was applied to quantify the importance of terms within the narratives. This approach identified high-frequency regulatory or financial terminology (e.g., "interest rate," "credit report") and distinguished them from generic or less significant terms.

3. **Dynamic LLM Analysis**: GPT-4 was employed to assess nuanced aspects such as tone, coherence, and the contextual impact of redactions. Prompts were refined iteratively, balancing specificity with the holistic strength of narratives.

<u>Example Prompt for Redactions</u>:

> *"Analyze the sophistication of this consumer complaint narrative. Consider how redacted text (e.g., 'XXXX') impacts specificity, clarity, and regulatory awareness. Adjust the sophistication score based on the narrative's ability to convey meaningful content despite redactions."*

<u>Validation and Refinement</u>

Validation was a critical aspect of the development process:

1. **Manual Review**: Subsets of narratives were manually reviewed to ensure alignment between human judgment and automated scores.

2. **Feedback Loops**: Adjustments were made to scoring weights and prompts based on observed discrepancies, particularly in edge cases with excessive redactions or vague phrasing.

3. **Comparison Across Tools**: Both Python and R implementations were tested on the dataset to ensure consistent outputs, enabling scalability for larger datasets.

Outputs and Applications

For each narrative, the methodology produced:

1. A **Composite Sophistication Score**: A numerical score (0–1) representing the overall quality of the narrative.

2. **Dimensional Breakdowns**: Individual scores for dimensions such as regulatory awareness, financial terminology, and clarity.

3. **Actionable Code**: Python and R scripts that automate the entire process, allowing scalability for analyzing large datasets.

Contributions

This methodology demonstrates the value of integrating regulatory insights, computational techniques, and AI-assisted collaboration:

1. **Framework for Textual Analysis**: Provides a scalable methodology for analyzing unstructured consumer data.

2. **Dynamic Treatment of Redactions**: Balances the privacy concerns of redacted text with the need for analytical rigor.

3. **TF-IDF-Enhanced Analysis**: Enables precise identification of critical regulatory and financial terminology.

4. **Actionable Tools for Researchers**: Python and R scripts enable researchers to replicate and extend this analysis on larger datasets.

By documenting this process, we hope to provide a roadmap for researchers seeking to combine traditional NLP techniques with LLM and TF-IDF capabilities, addressing complex textual challenges in regulatory and behavioral contexts. The methodology developed in this study can be adapted to other types of unstructured textual data, showcasing the versatility of LLMs in empirical research.

Methodology

This study examines consumer complaint narratives submitted to the Consumer Financial Protection Bureau (CFPB) to derive actionable insights into consumer behavior and resolution expectations. The analysis employed an iterative and collaborative approach, leveraging a Language Learning Model (LLM), specifically ChatGPT, to design novel metrics and analytical frameworks. The study combined text data with supplementary regulatory documents to contextualize the findings.

Data and Context

The primary dataset comprised anonymized consumer complaint narratives, where sensitive information was redacted with placeholders like `XXXX` in adherence to the CFPB Scrubbing Standards. These narratives provided the foundation for understanding consumer behaviors and resolution demands. To enhance contextual understanding, the CFPB Examination Manual was reviewed, offering insights into regulatory practices and escalation patterns within the financial ecosystem.

The study's objectives were outlined in an open-ended prompt to the LLM:

> "Finally, I am uploading the complaints data. I would love to hear your thoughts on how to examine it. Ultimately, I would like to use the complaints narrative to understand how consumers behave. So it would be helpful to come up with a measure. I have some ideas I could share."

This approach allowed the LLM to proactively contribute foundational ideas, enabling the development of metrics that aligned with both theoretical and practical considerations.

Metric Development

Through iterative interactions, the LLM suggested three key metrics:

Action-Oriented Score

- Quantifies actionable requests in narratives (e.g., "resolve," "refund").
- Captures the extent to which consumers explicitly demand solutions.

Resolution Expectation Metric

- Measures the clarity and specificity of resolution demands, including mentions of monetary amounts or timelines.
- Scaled to enhance differentiation across complaints.

Behavioral Indicators

- Captures consumer tendencies such as:
    1. **Escalation:** References to legal action or external reporting.
    2. **Frustration:** Linguistic markers of urgency or dissatisfaction.
    3. **Resolution Orientation:** Cooperative, solution-focused language.
- Combined into a Consumer Behavior Index with equal weighting for comprehensive analysis.

Prompts such as:

> "Yes, can you give me an example of a resolution expectation metric? Now think of the following—if there are other borrowers from a bank reading the complaints filed, which indicator/measure will help them in making a decision of whether to continue banking?"

guided the refinement of these metrics, ensuring practical relevance and stakeholder alignment.

Metric Computation

The metrics were calculated as follows:

- **Text Preprocessing:** Narratives were cleaned, tokenized, and normalized to handle placeholders like `XXXX` and focus on meaningful content.
- **Keyword-Based Analysis:** Linguistic features were extracted using targeted keyword lists for each metric.
- **Score Calculation:** Normalized keyword frequencies relative to narrative length ensured fairness across varying complaint lengths. Scores were scaled (e.g., multiplied by 100) to enhance interpretability.
- **Composite Indices:** Behavioral Indicators were aggregated into a single index using equal weights.

Role of the LLM

The LLM was integral at every stage of the research process:

1. **Proactive Metric Identification:** Suggested metrics based on the dataset and regulatory context without explicit instructions.
2. **Iterative Refinement:** Dynamically responded to user feedback, refining metrics and introducing scaling mechanisms.
3. **Scalability:** Proposed workflows for processing large datasets, ensuring feasibility for millions of rows.
4. **Contextual Analysis:** Synthesized insights from CFPB documents to align metrics with institutional and regulatory priorities.

Discussion

The integration of the LLM into the research process enhanced the methodology by introducing efficiency, scalability, and innovation. ChatGPT's contributions can be categorized as follows:

Metric Development

- The LLM proactively identified key analytical dimensions (e.g., action orientation, resolution expectation) based on the dataset and regulatory framework.
- Through iterative exchanges, the researcher refined these metrics, ensuring their practical application to stakeholders, such as borrowers evaluating financial institutions.

Automated Analysis

- The LLM automated tasks such as keyword extraction and normalization, reducing manual effort while maintaining analytical rigor.
- The scalability of the workflows proposed by the LLM ensured that millions of narratives could be processed efficiently.

Contextual Insights

- By synthesizing the CFPB Examination Manual and Scrubbing Standards, the LLM linked linguistic patterns in the narratives to systemic issues, such as escalation tendencies or resolution demands.

Research Innovation

- The LLM uncovered new dimensions for analysis, such as consumer persistence and sentiment trajectory, enabling a deeper understanding of behavioral patterns.
- Its dynamic interaction model facilitated a collaborative exploration of research questions, enriching the methodological framework.

Table OA.1: Effect of Complaint Action-Orientation Score on Deposits

This table presents the regression results of ChatGPT generated textual analysis of consumer complaint narratives assessing the effect of narratives' action-orientation score on deposits. *Action orientation* is the average action orientation score of complaint narratives for bank b in quarter t. The measure is a continuous variable from 0 to 10 quantifying the actionable requests in the narratives ("resolve" or "refund"). The measure captures the extent to which consumers explicitly demand solutions. Bank controls are lagged by one quarter and include liquidity ratio, capital ratio, ROA, and log(assets) as reported on quarterly Call Reports. Sample includes 79 CFPB supervised banks from 2015 Q1 to 2019 Q4. Bank controls are lagged by one quarter. Robust standard errors clustered by bank and quarter. All variables are defined in Appendix A. ***, **, and * represent statistical significance at the 1%, 5%, and 10% levels, respectively.

	Total Deposits	Insured Deposits	Uninsured Deposits
	(1)	(2)	(3)
Action score	0.015	0.008	0.023
	(0.014)	(0.021)	(0.015)
Bank size	1.025***	1.156***	0.900***
	(0.176)	(0.188)	(0.282)
Liquidity ratio	-1.962*	-2.365**	-2.571*
	(0.938)	(1.094)	(1.423)
NPL ratio	-1.879	-4.904	7.763
	(2.462)	(3.085)	(5.316)
ROA	-23.260	-36.571	32.057
	(30.673)	(41.049)	(34.546)
T1 capital ratio	-4.504	-5.382	-7.075*
	(2.607)	(3.153)	(3.997)
Bank FE	Yes	Yes	Yes
Quarter FE	Yes	Yes	Yes
Observations	1,547	1,547	1,547
Adjusted R^2	0.909	0.880	0.907

Table OA.2: Effect of Complaint Escalation Tendency on Deposits

This table presents the regression results of ChatGPT generated textual analysis of consumer complaint narratives assessing the effect of narratives' escalation tendency on deposits. *Escalation tendency* is the average escalation tendency of complaint narratives for bank b in quarter t. The measure is a continuous variable from 0 to 10 capturing references to legal action or external reporting. Bank controls are lagged by one quarter and include liquidity ratio, capital ratio, ROA, and log(assets) as reported on quarterly Call Reports. Sample includes 79 CFPB supervised banks from 2015 Q1 to 2019 Q4. Bank controls are lagged by one quarter. Robust standard errors clustered by bank and quarter. All variables are defined in Appendix A. ***, **, and * represent statistical significance at the 1%, 5%, and 10% levels, respectively.

	Total Deposits	Insured Deposits	Uninsured Deposits
	(1)	(2)	(3)
Escalation tendency	0.072	0.039	0.154
	(0.089)	(0.151)	(0.107)
Bank size	1.026***	1.156***	0.900***
	(0.176)	(0.187)	(0.282)
Liquidity ratio	-1.962*	-2.364**	-2.570*
	(0.938)	(1.094)	(1.424)
NPL ratio	-1.895	-4.913	7.736
	(2.472)	(3.095)	(5.307)
ROA	-23.200	-36.537	32.154
	(30.634)	(40.995)	(34.523)
T1 capital ratio	-4.504	-5.382	-7.073*
	(2.606)	(3.152)	(3.996)
Bank FE	Yes	Yes	Yes
Quarter FE	Yes	Yes	Yes
Observations	1,547	1,547	1,547
Adjusted R^2	0.909	0.880	0.907

Table OA.3: Effect of Complaint Frustration Urgency on Deposits

This table presents the regression results of ChatGPT generated textual analysis of consumer complaint narratives assessing the effect of narratives' frustration urgency on deposits. *Frustration urgency* is the average frustration urgency score of complaint narratives for bank b in quarter t. The measure is a continuous variable from 0 to 10 measuring the linguistic markers of urgency or dissatisfaction. Bank controls are lagged by one quarter and include liquidity ratio, capital ratio, ROA, and log(assets) as reported on quarterly Call Reports. Sample includes 79 CFPB supervised banks from 2015 Q1 to 2019 Q4. Bank controls are lagged by one quarter. Robust standard errors clustered by bank and quarter. All variables are defined in Appendix A. ***, **, and * represent statistical significance at the 1%, 5%, and 10% levels, respectively.

	Total Deposits	Insured Deposits	Uninsured Deposits
	(1)	(2)	(3)
Frustration urgency	0.029	0.251	0.020
	(0.078)	(0.246)	(0.118)
Bank size	1.026***	1.156***	0.900***
	(0.176)	(0.187)	(0.282)
Liquidity ratio	-1.962*	-2.364**	-2.570*
	(0.938)	(1.095)	(1.424)
NPL ratio	-1.892	-4.918	7.744
	(2.470)	(3.088)	(5.310)
ROA	-23.222	-36.688	32.132
	(30.651)	(41.036)	(34.530)
T1 capital ratio	-4.505	-5.385	-7.076*
	(2.607)	(3.154)	(3.997)
Bank FE	Yes	Yes	Yes
Quarter FE	Yes	Yes	Yes
Observations	1,547	1,547	1,547
Adjusted R^2	0.909	0.880	0.907