Learning About Other Depositors through Earnings Calls[⋆]

Sunjin Park[a,*], Simpson Zhang[b]

[a] *Hongik University,*
[b] *Cambridge Centre for AI in Medicine*

September 30, 2025

Abstract

We examine if there are valuable signals about depositor behavior released through banks' earnings calls. We use textual analysis on earnings call transcripts to measure the extent of discussions on deposit-related topics. Based on our metrics, we find that an increase in deposit discussions from earnings calls predicts uninsured deposit outflows but does not predict insured deposit flows. The interpretation is that uninsured depositors are responsive to news about their banks' other depositors. Further analysis indicates that the outflow reaction is stronger for banks with depositor characteristics related to more alertness or responsiveness.

Keywords: banking, deposit flows, uninsured deposits, textual analysis, earnings calls

[⋆]This work was supported by the Hongik University new faculty research support fund.

[*]Corresponding author.

Email addresses: sunjinpark@hongik.ac.kr (Sunjin Park), dr.simpson.zhang@gmail.com (Simpson Zhang)

1. Introduction

Predicting depositor behavior is an important but challenging problem for the business of banking. Banks, especially if they have market power, highly value this relatively cheap source of funding (Drechsler et al. (2017)) and use it to provide longer-term credit to the economy. However, deposits are unstable and hard to control, as they are inherently either short-term or have no specified maturities. As such, deposits pose a serious challenge for risk management since banks cannot fully control the future inflow or outflow of liabilities (Bolton et al. (2025)). These risks have shown themselves throughout banking history, including recently through the run behavior of uninsured depositors during Spring 2023.

A burgeoning literature on deposits has provided some insights into the factors that affect deposit balances. Oftentimes the rate environment has been found to play a key role. The recent regional banking crisis in the US highlighted how rate hikes, due to impairing long-term asset values, can cause depositors to awaken and run (Jiang et al. (2024)), similar in spirit to the classical bank run literature such as Diamond and Dybvig (1983). In recent decades, monetary policy has also affected depositor behavior through quantitative easing, with the expansion of reserves at the Federal Reserve shown to have resulted in flighty deposits (Blickle et al. (2024)).

At the same time, studies have shown that the complexity of depositor behavior makes the topic highly challenging. First, uninsured and insured depositors face significantly different incentives, as uninsured depositors must be much more careful of the solvency of their institution. The differences in behavior between uninsured and insured depositors have been well documented in papers such as Iyer and Puri (2012) and Chen et al. (2022). In addition, depositor behavior has been found to differ significantly between online and traditional banks. Erel et al. (2023) document that during periods of monetary tightening online banks experience deposit inflows while more traditional banks experience outflows. The failure of Silicon Valley Bank (SVB) in 2023 has been largely attributed to information spread on social media, highlighting the crucial role of information in precipitating bank runs (Cookson et al. (2024)).

We examine how depositors respond to informative signals generated through earnings calls. In particular we pay attention to deposit news to study how depositors learn and react to signals about what other depositors think of the bank's health, generating strategic complementarities as in Diamond and Dybvig (1983). A senior management of a bank can talk about deposit topics, such as the ongoing trend of its deposit funding or risks and opportunities in funding costs. Our premise is that these discussions about deposit topics ultimately provide positive or negative signals about how deposit customers currently perceive the bank's health. Since there are incentives for the bank to transparently communicate to avoid potentially adverse consequences of hiding material information, negative information revealed in earnings calls could serve as a signal to active depositors that most of the bank's depositors have negative perceptions of the bank. This then leads to negative feedback loops, as in the papers on bank runs, and the adverse perceptions among depositors creates large incentives to move funds elsewhere, resulting in significant deposit outflows.

We apply text methods on the transcripts of banks' earnings calls to measure the extent of discussions on topics related to deposits. To capture the complexity of discussions, we propose two types of text methods: one based on the count of keywords and another based on the degree of overlap with deposit-related topics created by Latent Dirichlet Allocation

(LDA). For the former, we ask ChatGPT to come up with a list of keywords that pertain to the topic of deposits. For the latter approach, we first conduct LDA, a topic modeling technique, and narrow down to a topic that is specifically about deposits. Then we construct metrics that are intended to measure how much of a given earnings call transcript overlaps with the keywords from the LDA output.

We then conduct a series of regressions to test if our constructed text-based metrics are able to explain deposit flows and balances. First, we find that more deposit discussions are followed by an outflow of uninsured deposits. In contrast, we do not find significant responses in insured deposit balances. This is consistent with our hypothesis that only uninsured depositors exhibit responsiveness to pertinent news about the health of their depositing institutions.

The earlier finding that deposits flow out in response to deposit discussions indicates that deposit discussion are associated with bank concerns. A priori, whether or not those discussions convey negative or positive information is ambiguous: bank management could also discuss deposits in an explicitly positive manner. For instance, they may say that the bank has experienced stable growth in deposit balances in spite of other financial performance concerns, in which case the anticipated stability in deposit funding is emphasized by management to mitigate these other concerns. We test this link by interacting our deposit discussion metric with either an overall sentiment score or the return on assets (ROA) as an indicator of adverse news. We find that the relationship between discussion score and deposit outflow is stronger for both negative sentiment and lower ROA. In another test, we also find that deposit discussions tend to be followed by an increase in non-performing loan balances. These findings suggest that the outflow response of uninsured accounts to deposit discussions is related to depositors' concerns. We can connect this finding with Chen et al. (2024), who document that uninsured depositor response to performance is stronger for banks with a larger extent of liquidity mismatch. Instead, our paper highlights the role of banks' discussions as a novel channel.

We then conduct several cross-sectional tests to examine for which banks the uninsured deposit outflows are stronger. We posit that the heterogeneity is driven by depositor characteristics, namely those related to depositor's alertness or responsiveness. Our channel is related to what extent depositors react to new information or signals released from conference calls. Hence, it is plausible that any difference in reaction is driven by the average characteristics of the depositor base that differs across banks.

Our findings collectively provide evidence that the response to discussions tends to be stronger for more alert- or responsive-related characteristics. We find banks that serve more urban depositors tend to exhibit higher deposit responsiveness. A possible reason is that urban customers tend to feel less attached to their depositing institutions and would be more willing to switch to other banks. Another finding is that banks serving higher expenditure population areas tend to show stronger responsiveness. An interpretation of this finding is that large spenders and rich customers maybe tied to well-performing businesses and have larger financial assets, causing these depositors to be more alert to financial news. Lastly, we also find suggestive evidence that banks serving more information sector-concentrated areas tend to show stronger responsiveness. We can interpret this finding as that customers in these areas tend to be better users of mobile banking, in which case, according to Koont et al. (2024), the propensity to exhibit reduced stickiness is higher.

We also test if deposit discussions are related to deposit pricing. According to Egan et al.

(2017), banks that are in distress compete aggressively for insured deposits by offering to pay higher rates, as part of an effort to make up the lost deposit funding from uninsured accounts. Martin et al. (2022) document that sophisticated investors tend to take advantage of increased deposit rates only up to the insured amount. Complementing their results, we find that the deposit discussion scores predict higher deposit rates for the insured time deposits. We also find that higher rates for these deposit accounts are then associated inflows of deposit balances, confirming that aggressive deposit pricing is used to make up for lost funding in uninsured deposits.

Literature Review

The study on banks' deposit flows can be tied to the panic-based run literature. Diamond and Dybvig (1983) highlight the strategic complementarities that result in self-fulfilling runs when depositors withdraw because they expect other depositors to withdraw. The theory of Goldstein and Pauzner (2005) provides the framework on how shocks to fundamentals are amplified by coordination failures. A recent work by Egan et al. (2017) explains the banking sector fragility through a structural model highlighting the behavior of insured and run-prone uninsured depositors. In their model, uninsured depositors are sensitive to the bank's financial distress and may run, while insured depositors do not exhibit sensitivity to it.

Some empirical papers shed light on the behavior surrounding bank runs in particular. Recently, Cookson et al. (2024) highlight the role of social media on the decline of stock prices and deposit runs in the case of SVB. In another paper by Cipriani et al. (2024), they use high-frequency interbank payments data to document empirical facts about the banking crisis in 2023.

There is also a growing literature providing various empirical evidences to understand depositors more broadly. Iyer and Puri (2012) document that deposit insurance reduces depositor runs and also show, for instance, that depositor runs are mitigated with longer bank-depositor relationship. The work by Benmelech et al. (2023) highlights the ongoing trend on banks with declining branch density and that, at the time of deteriorating health, banks experience uninsured deposit outflows. In the paper by Chen et al. (2022), uninsured depositor flows' sensitivity to performance is dependent on the banks' transparency. This finding is relevant for our study because we specifically pay attention to the information that is transmitted via earnings call meetings. At the same time, Jiang et al. (2022) find evidence that deposit inflows predict lower need for external capital market funding, reducing voluntary information disclosure. Hence, we can appreciate the importance of the interplay between deposit flow and information transmission.

Another strand of literature highlights interest rate risk in the context of deposit franchise value, and notable discussions on this topic can be found in DeMarzo et al. (2024) and Drechsler et al. (2021) and Emin et al. (2025). Specifically, the extent of the stickiness of deposits has become of a topic of interest as the US and several advanced economies experienced a steep increasing rate environment in 2022 through 2023, resulting in notable bank failures[1].

There is also a large literature, mostly based in the accounting field, that assesses firm

[1] See the publication by the Financial Stability Board: https://www.fsb.org/2024/10/depositor-behaviour-and-interest-rate-and-liquidity-risks-in-the-financial-system-lessons-from-the-march-2023-banking-turmoil/

discussions in earnings calls and the impact on firm stock returns and performance. While not focusing on banking specifically, these papers dissect various aspects of conference calls and show that even minor discrepancies can have significant impacts on subsequent stock returns. Several of these papers study the tone of a conference call, and they find that positive tones tend to be related to higher future stock returns. Price et al. (2012) finds that conference call tone is significantly associated with abnormal returns. De Amicis et al. (2021) also finds that stock market reactions are correlated with the sentiment of earnings calls. The paper by Amoozegar et al. (2020) assesses the impact of the presence of institutional investors on conference call tone. Short-term institutional investors are associated with a positive tone, while long-term institutional investors are associated with a negative tone. Call et al. (2024) finds that managers use of humor during conference calls leads to more positive stock returns, and they argue that humor helps mitigate the negative impact of bad news. Bochkay et al. (2020) finds that managers that use more extreme language during earnings calls will lead to a stronger subsequent stock price reaction and higher stock trading volumes. Chen et al. (2018) present contrasting evidence by finding that intraday stock prices respond more strongly to the analyst tone than to the manager tone in earnings calls.

Other papers have shown that the level of preparation and discussion present in conference calls can be significant. Lee (2016) finds that managers respond to questions using prepared scripts to avoid disclosing bad news. It finds that a lack of spontaneity is associated with negative abnormal stock returns and downgrades in analyst forecasts. Rennekamp et al. (2022) finds that earnings calls with more back and forth discussions between the firm and investors is associated with abnormal stock returns, providing evidence that such calls are more informative.

Notably, Cook et al. (2023) is a recent working paper that assesses the ability of LLMs to analyze bank earnings calls. They use LLMs to assess the topics discussed, sentiment, temporal orientation, and vagueness of the earnings calls. The paper does not focus specifically on deposit discussions, but it does find that discussions of deposits are the most commonly discussed among all topics in these earnings calls. Their sentiment analysis shows a higher sentiment during the presentation section than the Q&A section of the earnings call. Discussions of deposits are found to be of generally high sentiment. Our paper expands on this work by considering a much longer time series, and connecting with data on deposit flows to show the relationship between deposit discussions and actual deposit flows.

2. Method

2.1. Data Description

For our study, we download banks' earnings call transcripts available from `seekingalpha.com` (Seeking Alpha). Earnings release meetings are typically held quarterly, and Seeking Alpha makes available the transcripts for many of the large, publicly traded banks in the US. Since transcripts are not widely available from Seeking Alpha for meetings held in 2007 or earlier, we restrict our sample period to 2008 and onward and include up to the end of 2024. Textual analysis is conducted for each transcript, which starts from bank management's earnings release presentation and then moves onto a question-and-answer section responding to equity analysts' questions.

These transcripts are then merged with the Call Reports data that is publicly available from FFIEC (`cdr.ffiec.gov`). For the most part, based on bank names we hand-match banks with call report data to those for which we have earnings call transcripts. Since banks often have layers of entities, which also change over time, we make use of the Relationships table available from the FFIEC website (ffiegov/npw/FinancialReport/DataDownload), which lists parent and child RSSD IDs with specified relationship dates. From the resulting merge, we select a single RSSD ID for each Seeking Alpha bank based on whichever RSSD ID has the largest total deposits at a given point in time.

Call Reports include a quarterly snapshot of balance sheet items, income statement items as well as other measures relevant for regulatory purposes. Uninsured deposits are computed as the total deposits minus insured deposits. A variety of performance-related items such as return on assets, net interest margins defined as net interest income divided by total assets, loan loss provision ratio defined as the credit loss provision divided by total loans, pre-provision operating profits defined as the sum of net interest income and net non-interest income divided by total loans, and non-performing loan balances. Capital structure-related items that we retrieve include: debt-to-equity, loan-to-deposit ratio, and core deposits divided by total assets.

We also make use of the Summary of Deposits from the FDIC, from which we retrieve the physical distribution of deposits at the branch-level. Finally, several types of economic data are retrieved from the Bureau of Economic Analysis website, including branch-level indicators of urban, state-level measures of personal consumption expenditure and industry contribution to the state real GDP.

2.2. Construction of Text-Based Metrics

In this section, we present two approaches for measuring the extent of deposit-related discussions by banks' management. The first measure will be based on the count of keywords, and the second measure will build on top of topic modeling.

2.2.1. Keyword-Based Deposit Measure

Our primary method of measuring the extent of deposit discussions is based on the frequency of relevant keywords. The advantage of this approach is that it is straightforward to understand what the metric captures. After coming up with a library of keywords, which we describe shortly, we then count the number of times c_l^i each keyword l appears in a given transcript i, repeat this for each keyword in our library L, and sum over all of those keywords.

$$wordcount_{i,t} = \sum_{l \in L} c_l^i \tag{1}$$

We also construct a scaled version of $wordcount_{i,t}$ by dividing it by the total number of words $totalwords_{i,t}$ in the given transcript i

$$wordcountintensity_{i,t} = \frac{wordcount_{i,t}}{totalwords_{i,t}} \tag{2}$$

Some banks inherently have more topics to cover because of the wide range of businesses they are involved with, so scaling by the total number of words helps measure the "density" of the discussion surrounding deposit-related topics. However, the absolute quantity of deposit discussion is also relevant because irrespective of the total length of the conference meeting, having more information on deposit behavior provides more information to process.

The obvious challenge of this approach is to come up with a good library of deposit-related words. We ask ChatGPT (o3 model) for a list of deposit-related keywords by inputting the following prompt:

I am trying to do text analysis on bank earnings call transcripts. For each transcript, I want to measure how much discussion there is on the topic of deposits. Instead of just searching for the word "deposit", I would like to search for keywords that are related to topics on deposits. Please give me a library (word list) of keywords that pertain to topics related to deposits.

We then count the number of sentences that contain each word from this keyword list. The response from ChatGPT on the list of deposit-related words can be found in Table 2. Although ChatGPT conveniently maps each keyword into a category of deposit topics, we simply aggregate across categories. Also, in the implementation of our calculation we simply replace all bigrams (or trigrams) containing the string "deposit" by the unigram "deposit", which means that we just search for the word "deposit" instead of searching for the associated bigrams (trigrams). Also when searching for bigrams (trigrams), we count as a match if the two (three) words all appear within the same sentence. We discuss the outputs of the word count approach further in Section 2.3.

2.2.2. LDA-Based Deposit Measure

In addition to the method using keyword counts, we also compute a metric based on a topic modeling method. The extent of overlap between a given transcript and our derived deposit topics will measure the extent of deposit discussions. We describe the steps below.

A topic modeling method called Latent Dirichlet Allocation (LDA) will be used to generate topics. LDA is a Bayesian model, in which a topic is modeled as a mixture over an underlying set of topic probabilities (Blei et al. (2003)). This model uses a bag-of-words approach, which means the input document is a large list of words without consideration of order. We supply a large number of “documents" as input after dissecting a long text of corpora separated into small pieces. The model then attempts to assign topic probabilities for each “document", for which topics are formulated as a group of keywords, each of which has differing importance.

We make use of earnings call transcripts since they include deposit topics, among other topics. In other words, we run LDA on the transcripts to extract topics and specifically restrict to the topic that is related to deposits. For computational efficiency, a random sample of 500 transcripts will be drawn, and each will be dissected into small pieces as inputs. While paragraphs seem a natural choice as input documents, there is not a clear delineation of paragraphs in the transcripts, so we instead cut off after 10 consecutive sentences and refer to each ten-sentence as a "document." After dissection, the collection of 500 random transcripts will help the model determine the topics.

LDA starts with a set of N topics and then iteratively updates the word assignment based on an algorithm until the process stabilizes.[2] Initially, words are randomly assigned. We specify the number of topics N to be 20 and apply the collapsed Gibbs sampling for the estimation. LDA continues reassignments for a set number of iterations. The final output of the LDA consists of N topics, each of which has probabilities assigned to each keyword, describing the corresponding topic is about.

After 20 topics are produced as outputs, we will examine which of those are related to deposits. For each topic, a list of words with the associated probabilities is produced, allowing us to determine if for any given topic the word "deposit" is one of the top-probability words.

[2]The input documents need to be cleaned. We remove common words that do not help identify topics (e.g., “will", “can", "also", “see") and also stem words to combine tenses.

It turns out only one of the 20 topics has the word deposit as a high-priority word, so we proceed with that specific list of words along with the associated probabilities. Further discussion on the interpretation of that topic can be found in Section 2.3 when we discuss the textual analysis outputs.

Once the deposit topic is formed, we move onto the second stage, that is, to measure the extent to which the text in each transcript overlaps with that topic. The main objective is to create a measure that is higher when a transcript covers keywords in those topics with a higher frequency. We use the method proposed in Zhang (2023) to measure deposit discussion metrics. Given the list of words R from topic j produced from the first stage as the output topic on deposits, we count the number of times that word k from R appears in transcript i and denote that as w_k^i. This w_k^i is then weighted by the probability β_k of the word k representing topic j, for which higher β_k indicates that the word k is more important for topic j. The length score can be written as follows

$$length_{i,t} = \sum_{k \in R} \beta_k \times w_k^i \tag{3}$$

Similarly as before, we scale the length score by dividing it by the total number of words in $totalwords_{i,t}$ and refer to it as the match score:

$$match_{i,t} = \frac{length_{i,t}}{totalwords_{i,t}} \tag{4}$$

2.3. Textual Analysis Outputs

We first look into intermediate and final outputs priduced by our text methods. The deposit-related keywords produced by ChatGPT are shown in Figure 2. The output classifies keywords into one of six categories: (1) Competitive & strategic language, (2) Deposit dynamics & metrics, (3) Deposit products & account types, (4) Regulatory & policy references, (5) Fee & revenue items tied to deposits, and (6) Funding & liquidity terminology. Hence, we can see that the deposit-related topics are intended to cover a variety of topics, such as account balance trends, strategic decisions like promotional-type accounts, projection on the rate environment and deposit pricing.

We then examine how the extent of deposit discussions within each categories vary over time. We pay attention to instances of significant variations over time. We note that categories covering funding and liquidity, trend in deposit dynamics, and product types experienced a substantial increase in 2017 subsequently dropped and then spiked up during 2022 to 2023 to an even larger extent. From a visual look it appears that these sub-topics increase during times of rate hikes.

Another observation of interest is the time-series for regulatory and policy references. Unlike any of the other categories, this specific category is the only one to experience an increase in frequency during 2008 to 2012. One possible reason is that at that time there were concerns with the instability of deposits when many banks experienced performance issues. According to a Notice of Proposed Rulemaking in 2008, "[a] number of costly institution failures, including some recent failures, have experienced rapid asset growth before failure and have funded this growth through brokered deposits."[3]

As mentioned earlier, we highlight that the resulting deposit metrics based on this library do not consider the categorization and simply aggregate across keywords. Moreover, for bigrams or trigrams with the string "deposit", we instead simply search for the word "deposit." The second type of text metrics uses LDA topic modeling in the first stage. Based on specifying a total of N=20 number of topics, our observation shows that Topic 12 happens to have the word "deposit" as the top-priority word. A word cloud of Topic 12 is shown in Figure 2.

Since it is difficult to visually tell what the collection of words indicate, we feed in this output and ask ChatGPT for its interpretation of this word cloud. The response is provided in Figure 3. One of the frequent sub-topics is the discussion on the rate environment. Recent or anticipated Fed actions and the resulting yield curve shape is an important topic from a funding perspective, and deposit cost is a relevant angle. The quality and composition of deposit accounts are also an important sub-topic, and it indicates how banks monitor the more stable and sticky core deposits separately from the more volatile deposits like brokered deposits and wholesale deposits. Deposit repricing mechanism is also an important sub-topic, indicating management's strategies on how to respond when the reference rate goes up or down.

Lastly, we visually examine the time-series of the different specifications of deposit discussion metrics. Throughout the empirical results section, we examine four specifications in parallel: the keyword-based metric and the LDA-based metric, each with the raw version

[3]https://www.fdigov/resources/regulations/federal-register-publications/2008/08propose1016.pdf

and the scaled version after dividing by the total number of words. We demean each metric and divide by the standard deviation, so that the standardized metric will be used for all subsequent regressions. For visualization purpose only, we average each metric across banks and present the resulting time-series in Figure 4.

Comparing across the four specifications, we observe that there is a noticeable commonality for the latter two thirds of the sample period. We find that starting from 2014 and onwards, the four specifications result in highly correlated time-series that appear to have some relationship with rate hike cycles. Another notable observation is that there is noticeable discrepancy across the specifications prior to 2014. The earlier period was presumably heavily impacted by the Global Financial Crisis, during which the struggle within the US financial sector was much more adverse than the latter part of the sample period.

3. Hypothesis Development

We are primarily interested in the informational content of the deposit discussions that occur during banks' earnings calls. For the text-based method, the value of the metric would be high if a bank devotes a substantial amount of text to discussing deposit-related topics. A higher extent of deposit discussions indicate that the management of a bank has a lot to update about its depositors, such as whether deposit accounts of certain types are increasing or decreasing, what kinds of strategic business decisions are made in relation to deposit funding, and how the business is currently handling or in the future plans to handle deposit costs. Since these updates and plans are inherently tied to how the bank's customers are thinking about the bank, we interpret deposit discussions as signals about what the bank's depositors perceive of the bank's conditions. If management talks about dealing with deposit outflows from certain regions or account types, that signals negative perception among those departing depositors. If management needs to aggressively raise rates to attract deposits, this could also signal negative perception because it could be a strategic decision to attract fleeing depositors despite the increase in costs. Once these are communicated externally, depositors learn from these news that the other depositors are thinking negatively about their depositing institutions. Strategic complementarities among depositors' learning from the behavior of other depositors result in additional deposit withdrawals. Hence, overall we would expect more deposit discussions to lead to more deposit outflows.

We posit that the relation between deposit flow and deposit discussions depends on whether or not deposits are insured. Customers with accounts that are insured should be insensitive to short-term news about the financial health of their bank. Especially in the US, the deposit insurance system promises account holders the ability to seamlessly regain access to insured balances in case of bank failure, thus creating a sense of stability for insured depositors even if the bank faces financial distress.

On the other hand, account holders with uninsured balances have a stronger incentive to monitor the viability of their depositing institutions. If there are signs that the health of their bank is deteriorating, uninsured depositors would be motivated to move their funds elsewhere.

It is plausible that a large fraction of the customer base is probably not following news about the short-term fluctuations in banks' health since the risk of incurring a loss due to bank failure is not a day-to-day concern. Moreover, earnings call meetings are typically considered to be more targeted to equity investors, acting as a common channel for banks to provide updates on ongoing financials to investors. The channel has to be that, based on the presentation and back-and-forth dialogue between bank's management and call participants, the customers of the bank would either directly listen to updates or indirectly receive information through news outlets and respond if needed. The key mechanism is that valuable information is communicated in these meetings, and deposit customers can readily access that information and then respond by reallocating their funds.

First, we test for the following main hypothesis regarding the behavior of uninsured deposits.

Hypothesis 1: Uninsured deposits flow out in response to substantial deposit discussions occurring in banks' earnings calls.

In contrast, deposit balances of insured accounts would be relatively insensitive to deposit discussions from earnings calls. Although uninsured deposit balances may flow out because of adverse news, we postulate that insured deposit accounts would not flow out from the same adverse news because there is no imminent risk of losing access to their funds.

When bank management discusses deposits, the eventual effect on deposit flow is ambiguous. One possibility is that banks tend to provide adverse updates on how certain types of deposits are experiencing sluggish growth, in which case the prediction is that uninsured depositors would flow out over concerns about bank health. However, the opposite direction is also possible. It may be that banks actively discuss positive topics such as high growth in their deposit accounts with the intention to please investors and attract positive attention from financial markets. If such a strategy is effective, then deposit balances, including uninsured deposits, may grow because banks could attract new customers, either in retail or commercial, who want to access financial services from a well-performing bank. Our prior, however, is the former, that banks tend to feature negative deposit discussions that lead to a deposit outflow. We conjecture that the decision to flee in response to adverse news is of bigger importance than responding to positive news.

The next set of hypotheses seeks to confirm whether or not the decision to flee is really due to learning about adverse news. We first examine if the flow response is stronger when negative news is more adverse. We specifically test two different specifications: one based on the tone of the earnings call and another one based on the latest performance from call reports. Potential performance issues will become salient when the tone of the call is negative, in which case a depositor would feel more pressured to react. The latter approach captures adverse news by using a standard metric of performance, namely the return on assets. Either way, we measure if the deposit flow response to deposit talks is stronger due to increased concerns about bank health.

Hypothesis 2a: The deposit flow response is stronger when the tone of the call or the performance indicator indicates adverse news.

We extend this line of analysis by testing if our deposit discussions predict real outcomes. If it is true that deposit discussions are related to the decision to move funds elsewhere, then presumably a larger extent of such discussions make adverse outcomes more likely. For instance, we examine the growth in non-performing loans. In case there is evidence that higher deposit discussion scores predict adverse performance in subsequent quarters, we can confirm that our earlier finding that depositor outflow is driven by the increased likelihood of upcoming performance deterioration.

Hypothesis 2b: More substantial deposit discussions are followed by worsening economic conditions for the bank.

We then conduct a series of additional tests to shed light on any heterogeneous response across the cross-section of banks. Specifically, we conjecture that the differences in response to deposit discussions are related to the characteristics of the depositor base. One of the hypotheses is that the deposit flow response is stronger for banks with a more urban customer base. Our reasoning is that urban customers are less tied to the local banking relationship and are more likely to quickly switch from one banking relationship to another. In another test, we hypothesize that the depositor flow response is stronger for banks located in population with large consumption expenditures. Our prior is that large spenders are more likely to be “alert" to financial news. Lastly, we examine if having more tech-savvy customers is associated with increased likelihood of shifting around funds frequently. Our reasoning is based on the finding by Koont et al. (2024) that digital banking has made it more likely for depositors to walk away. We proxy for this by measuring to what extent the local economy is focused on the information sector.

Finally, we also examine the extent of the bank's revenue mix. Our hypothesis is that depositors of banks that engage in a variety of financial services other than traditional interest income tend to be more responsive. We posit such banks are also likely to attract deposits from the related businesses. It could be that these deposit account holders are better-informed about the health of the depositing institutions and also more active in managing the funds in their deposit accounts. Alternatively, even if the deposit account holders are not affiliated with those businesses, we would find similar results as long as such banks attract deposit holders that are at least better-informed about the broader financial services sector, including trading and M&A activities.

Hypothesis 3: Depositor characteristics, such as alertness or responsiveness, affect depositor's response to bank discussions.

In the final section, we examine the pricing aspect of deposits. Our overall line of hypothesis testing is presumably also related to banks' decision to strategically price deposit rates. Reading over a sample of earnings call transcripts, we frequently encountered discussions on the bank management views on how they intend to price deposit accounts in the near-term future, often called the deposit beta. Banks may be discussing deposit rates in the context of how they contribute to the latest trends on net interest income, or in the context of how competitively a bank may position itself in a market during times of rising interest rate environment. Hence it is natural to test whether the depositor flow response is related to changes in deposit rates.

More specifically, we build on the finding of Martin et al. (2022) that financially distressed banks aggressively raise deposit rates to term deposits, thereby attracting funding from more sophisticated depositors who intend to enjoy the favorable pricing, yet only up to the FDIC insured limit. Accordingly, we examine if our deposit discussions are also followed by higher deposit rates for such account types. Our premise is that if more deposit discussion is associated with concerns about the healthiness of the bank, then those banks may choose to attract deposit funding by offering more favorable deposit rates for certain deposit account types.

Hypothesis 4: Banks raise insured time deposit rates after substantial deposit discussions.

To verify if this funding strategy is indeed effective, we test if higher deposit rates are followed by higher deposit inflows. Again, we conduct our analysis at the account type level because the strategy is intended to pursue a specific type of account holders, namely sophisticated investors.

Lastly, we conduct additional analysis to provide a more comprehensive understanding of depositor responses. One regression analysis we highlight is whether the flow response is stronger depending on the level of uninsured deposit funding. According to Chen et al. (2024), larger liquidity mismatch, which can be measured using the fraction uninsured, leads to higher bank fragility. While our set of hypotheses does not necessarily tie to liquidity mismatch specifically, we do analyze how uninsured depositor responsiveness is related to deposit discussions so that we can better tie to the existing literature.

Regression Specification

In general, we run panel regressions with fixed effects and control variables as follows:

$$y_{b,t+1} = depositdiscussion_{b,t} + X_{b,t} + \gamma_b + \gamma_t + \varepsilon_{b,t+1} \tag{5}$$

The main independent variable of interest is the deposit discussion score computed using one of our proposed approaches. As discussed above, one of the approaches is based on a keyword count of deposit-related words. The other approach is an overlap metric in the second stage after initial LDA topic analysis on a set of deposit-related documents. The dependent variable of interest in most of these regressions is the deposit flow, measured as the change in deposit balances scaled by total assets. The deposit balance can be uninsured balances, insured balances, or total deposit balances. Uninsured deposit flow observed from t to $t + 1$, for instance, is measured as the quarterly change in uninsured deposit balance from t to $t + 1$ divided by the total assets observed at t.

Pertinent controls are included on the right side of the regression equation. We include bank fixed effects which allow for time-invariant differences across banks to be absorbed. Quarterly time fixed effects are also included, controlling for economy-wide differences across time. We also include bank-specific profitability measures and funding-related balance sheet measures. The inclusion of these variables is intended to control for typical quantitative variables that become available from bank call reports. In the absence of other communications, investors, depositors, and other stakeholders can learn about the financial health of banks through the quarterly publication of these regulatory reports.

4. Empirical Results

In Section 2.2 we computed metrics on the extent of discussions on deposit-related topics during earnings calls. In this section, we empirically examine if these discussions contain useful information about deposit flows.

4.1. Main Regressions

Our main regression analysis is to test whether our measures of banks' deposit discussions during earnings calls predict deposit flows. For all of the regression tests in our work, we conduct four different specifications of deposit discussions. The first is the keyword count metric based on counting the number of occurrences within the transcript of keywords in the deposit-related library. This metric is denoted as *depositcnt* in the subsequent tables. The scaled version of it is denoted as *depositscaled*. The third version is the length score, which measures the extent of overlap with the deposit topic keywords from LDA and is denoted as *depositlength*. Finally, the scaled version of it is denoted as *depositmatch*. For all of the discussion metrics, we standardize each so that it has mean zero and unit standard deviation.

Table 3 presents the predictability in three ways: the next period change in uninsured deposits from t to $t + 1$; the contemporaneous change from $t - 1$ to t; and the two-period ahead from $t+1$ to $t+2$. The first four columns indicate that our deposit discussion metrics load negatively in predicting the next period change with statistical significance in three of the four text metric specifications. The scaled keyword count metric *depositscaled* also loads negatively but with slightly weak t-statistics. The next four columns show the contemporaneous relationship. We can see that the relationship is generally negative with statistical significance only with the keyword count specification. Finally, the last four columns show the results with two-period ahead uninsured deposit growth rates as the dependent variables. Here, the discussion scores load negatively with statistical significance for two of the deposit talk specifications.

Collectively, the results from this table suggest that when banks discuss more deposit-related topics their uninsured deposit balances tend to go down. This relationship is present across different time stamps for deposit flows. The predictive relationship seems to be more robust compared to the contemporaneous relationship, which suggests that there is a lagged response in depositor behavior after when the discussion takes place. This observation points to the explanation that it takes time for depositors to fully react to the news generated in earnings calls. Based on this finding, in many of the subsequent analysis we conduct regressions with the prediction of the $t+1$ changes.

We would highlight that these regression results are robust after including relevant control variables. We include a number of profitability measures, such as return on equity, net interest margin, credit loss provision scaled by total loans, and pre-provision operating profit scaled by total loans. We also include a number of metrics on leverage and funding-related metrics from the balance sheet, including debt-to-equity ratio, loan-to-deposit ratio, and core deposit balance scaled by total assets. Our regression specification should help address plausible channels of performance and funding-related standard metrics driving deposit flows. Also included is the trend in percent changes in total deposits over the past 4 quarters. This helps capture depositors responding to past deposit balance trends. Finally, the bank and time fixed effects control for time-invariant bank heterogeneity as well as the overall economy-wide components of uninsured deposit flows.

Hence, our tests allow us to argue that, even after controlling for the standard dataset variables that become publicly available through the regulatory filings, the textual measures in various specifications reliably predict uninsured deposit flows. This finding supports the argument that the deposit conversations occurring in these earnings calls are indeed important. Uninsured depositors, specifically, do seem to respond to the information generated from these calls.

We then repeat the regression replacing the dependent variable with insured deposit growth instead. The results are shown in the first four columns of Table 4. We find that our deposit discussion metrics do not negatively predict insured deposit flows. If anything, the deposit scaled metric, which is based on keyword count, seems to predict uninsured deposit flows positively. We will later revisit this finding when we examine the deposit pricing channel. When we use the total deposit change as the dependent variable, the results are not statistically robust as shown on the latter four columns of the same table. The negative loading on deposit length score is statistically significant, however.

The finding on insured depositor behavior suggests that there is clearly a distinct response between insured depositors and uninsured depositors. Depositors with insured accounts should not be responsive to short-term fluctuations about their banks' financial conditions because bank failures do not pose any risk to these customers. In contrast, depositors with funds that are not fully insured may be concerned about potential bank failures. Since discussions only affect the outflow of uninsured deposit balances but not insured balances, we can plausibly suspect that whatever content we are capturing through our text metrics concerns uninsured depositors but not insured depositors.

A priori, it is unclear if our deposit metrics capture negative news. Our keyword based metrics *depositcnt* and *depositscaled* record higher values if there are many sentences with the words in Table 2, but these keywords do not have negative connotation. The word clouds in Figure 2, which highlight the main keywords of the LDA analysis also do not visibly exhibit any negative connotation.

We manually checked for sentences with the word deposit to get a sense of whether deposit topics were described in a concerning context. Based on randomly checking a few, we could not unambiguously tell if the bank was struggling just from narrowly reading the sentences. Deposits are often discussed in terms of deposit pricing as rate changes or is expected to change. Deposits are also discussed when explaining and comparing the trends of customer balances across different types.

Accordingly, in the subsequent test we repeat our main regressions, but we interact our deposit discussion score with a sentiment variable for the earnings call. The sentiment score of an earnings call transcript is measured using the standard Loughran-McDonald sentiment library that is specifically tailored to accurately measure the sentiment of financial words. We then construct an indicator variable for negative sentiment that scores one if the particular transcript scores below the median sentiment score on that year-quarter and zero otherwise.

As shown in Table 5 we find that there is strong evidence that, when the overall tone of a transcript is low, a higher deposit discussion score predicts stronger uninsured deposit outflows the next quarter. The results are statistically significant for all four specifications of deposit metrics. This finding strengthens our earlier argument that more deposit discussions predict outflows out of concerns because the response of the uninsured depositors is stronger with a more negative tone in the overall discussions.

In Table 6, we repeat the analysis but replacing the sentiment variable with return on assets (ROA). The intention is to examine if there is evidence of a stronger deposit flow response when the actually recorded bank performance is poor, as measured by ROA. We find from the results that lower ROA is associated with stronger flow response when we use the keyword count metrics. While the statistical significance is absent with the LDA-based metrics, directionally the results are consistent with the sentiment results that adverse news amplify the uninsured depositors' response to the discussions.

We then extend the analysis by testing if our deposit discussion metrics are capturing information about future adverse financial conditions. In the first four columns of Table 7 we regress the subsequent logged change in non-performing loans onto our deposit metrics. We see that higher deposit discussions seem to predict an increase in non-performing loans next quarter yet with generally weak statistical significance. In the next four columns, we replace the dependent variables with the two-period-ahead non-performing loan balances. We find that the results are stronger here than the one-period-ahead. Collectively, these tests indicate that higher deposit discussions are followed by an increase in non-performing loans, again corroborating our argument that concerns about adverse conditions seem to be driving our results.

4.2. Characteristics of Banks' Average Depositors

In this section, we examine if our finding on the relationship between deposit discussion metrics and future uninsured deposit flowa is stronger for certain banks. We examine this heterogeneity across banks by repeating our main predictive regressions but interacting the deposit talk metrics with another variable of interest. As explained in the hypothesis section, we test interactions with variables that can help differentiate the characteristics of each bank's average depositor.

We first consider a set of variables that are dependent on the geographic footprint of a given bank. The first variable of interest is a bank's extent of urbanness. From FDIC's summary of deposits (SOD), we retrieve data on the branch-level deposits for each bank. Then using SOD's indicator flag for a core urban area of 50,000 or more of population (METROBR), we compute what fraction of the bank's total deposits is located in urban areas. The intention of the urbanness variable is to examine if depositor behavior from urban banks is different from that from non-urban banks.

The first four columns in Table 8 indicate that there is evidence that urban banks exhibit higher sensitivity of uninsured deposit flows to our deposit metrics. The coefficients in front of the interaction between *urban* and each of our deposit discussion scores is negative with statistical significance in three of the four specifications. For an urban-focused bank, uninsured depositors flow out more when deposit talks increase. Our interpretation is that customers who deposit more at an urban-focused bank are more responsive to ongoing news about the healthiness of their depositing institutions. One possibility is that urban customers are more "alert", meaning depositors may shift around funds a lot when news spreads out and exhibit greater sensitivity to it. Another interpretation is that urban customers feel less attached to particular banks and are more inclined to abandon existing banking relationship and explore opening new accounts.

The next variable we consider is the average personal consumption expenditure (PCE) of the bank's depositor base. From the Bureau of Economic Analysis (BEA), we retrieve state-level PCE per capita and then for each bank we compute the weighted average of state PCEs, weighted by the bank's deposits across states. The value of this variable scores higher if a bank has a heavier presence in states with high expenditures. Our reasoning is that customers with a high amount of expenditures are more "alert" to financial news. Higher expenditures likely indicate having customers who are tied to well-performing businesses and who likely hold large investment assets. These customers are likely to pay attention to financial news, including any material news about their depositing institutions.

The fifth through the eighth columns of Table 8 show the results with interacting deposit metrics with average PCE of the bank's customers. We observe that the flow response is stronger when a bank's average depositor has higher PCE, and the statistical significance is strong for three out of four specifications. In unreported tables, we also find similar results when using local real GDP as the interacted variable. We also repeat with personal income, and the results are similar with statistically significant coefficients for two out of four deposit metric specifications.

The next variable of interest is the contribution of the "information" sector to the state economy. Again from the BEA, we retrieve data on the contribution of the information sector to the percent change in state-level GDP. Then, based on each bank's geographic distribution of deposits across states, for each bank we compute the weighted average of the information sector concentration. A bank with heavy presence in, say, California or the state of Washington would score high on this variable because the contribution of the information sector to the state's GDP tends to be higher for these states. The motivation for testing with this variable is that customers residing in states with strong information sectors may tend to exhibit higher tendency to utilize mobile banking. A possible implication is that customers may be more inclined to shift around their funds since digital banking has made it more likely for depositors to walk away from current banks when the situation arises. Alternatively, it could mean that information sectors generally attract financially more alert people, who show more responsiveness to financial news that they perceive as impacting their banking services.

From the last four columns of Table 8, we can see that banks located in information sector-heavy states tend to show greater sensitivity of uninsured deposit flows to our deposit metrics. The interaction coefficient is only statistically significant for *depositmatch*, while the other three specifications give marginally insignificant estimates. Nonetheless, we confirm the takeaway is consistent with our prior that information sector-focused locations tend to be more reactive to deposit discussions.

In another exercise, we explore cross-sectional differences in terms of bank characteristic. Specifically, we test whether banks with a higher focus in noninterest income experience different responses to deposit talks. The motivation for this angle is that pursuing revenues other than the traditional interest income could be associated with attracting deposit accounts from businesses that provide noninterest income. This is because it is natural for the bank to try to offer a combination of banking services including deposits, which could offer convenience from the client's perspective. A possibility is that these depositors, who also provide noninterest income to the bank, may be more responsive to adverse news because these clients would be relatively better-informed of adverse financial conditions.

Table 9 shows the results from introducing interactions between the *revmix* variable and desposit metrics. The *revmix*, here, is defined as noninterest revenue divided by interest revenue.[4] We find that the deposit outflow from an increase in the deposit keyword count is exacerbated with higher levels of the *revmix* variable. This means that uninsured depositors are more responsive to deposit news if the bank is characterized as having a higher pursuit of noninterest income.

4.3. Deposit Pricing

We first begin this analysis by controlling for deposit rates. In Table 10 the first four columns indicate that controlling for the overall deposit rates on the right hand side do not weaken the predictive ability of deposit discussions and that the deposit rates do not load significantly. Interacting deposit rates with our deposit discussion metric also indicates that the overall deposit rates do not seem to explain deposit flows much. This means that, at the overall level, banks' choices on deposit rates are not very much related to the response of uninsured depositors to banks' discussions on deposit topics.

Next, we examine if deposit discussions also explain future deposit pricing. As is clear in Figure 2 as well as ChatGPT's summary of the topic in Figure 3, deposit pricing is a frequent topic of interest in earnings calls. The rates that banks pay to their deposit customers are an important component of funding costs, in particular affecting banks' net interest margin. At the same time, banks carefully design deposit pricing strategies to keep their expenses manageable while maintaining their customer base. Sometimes, these strategies include

[4] We adopt the definition of Bushman et al. (2016), who find that competition results in banks to seek alternative sources of revenue.

management pursuit of particular customer segments to expand market share.

The finding of Martin et al. (2022), building on the theory of Egan et al. (2017), provides us guidance on the role of deposit rates, especially during times of financial distress. Those authors find that when banks are in difficult conditions, banks try to compensate for the outflow of uninsured deposits by raising deposit rates to attract insured depositors. In the study of Martin et al. (2022), the newly attracted deposits of failing banks were mostly insured term deposits just under the FDIC insurance limit. These accounts are most likely made up of sophisticated investors who are actively searching for the highest paying accounts but with limited risk. In essence, these banks aggressively raised insured term deposits as part of their funding strategy in difficult times.

Accordingly, we restrict to insured time deposits for this analysis. We regress the next-period deposit rates on insured time deposit accounts onto our deposit discussion metrics. We proxy for deposit rates by retrieving the quarterly interest expense on insured time deposits and dividing it by the balance of insured time deposit accounts.

The results are shown in the first four columns of Table 11. There is evidence that, when we use the keyword count specifications, higher deposit discussions lead to significantly higher deposit rates for insured time deposits. When bank management heavily discusses deposit topics, either directly by talking about deposit pricing strategies or talking about other deposit-related topics in general, there is an increased likelihood that banks may actively seek to offer high deposit rates to attract funding from rate-seeking depositors. As shown in the last column of the table, an increase in deposit rates for insured time deposits predicts more growth in the balance of time deposit accounts.

These pieces of evidence again suggest that deposit discussions are likely associated with deposit funding concerns. Specifically, as uninsured deposits flow out, banks may seek funding from other conceivably more active depositors by offering attractive rates.

4.4. Additional Analysis

In this section we conduct a series of additional tests to further test and take advantage of the empirical setup.

4.4.1. Robustness

We first conduct a robustness test that exploits the date of earnings calls. The results in Section 4.1 suggest that it takes time for depositor behavior to fully respond to news transmitted through earnings calls. We note that the as-of-date filing for the call reports is uniform across banks, while the earnings call dates vary across banks. Hence we conjecture that the impact of the news released from earnings calls is more fully realized for banks that have call dates earlier within the quarter than for those that have call dates later within the quarter. Our empirical design is to calculate the number of days from the call date until the end of the quarter for each bank and to create a dummy variable that equals one if this variable is above the median for a given year-quarter and zero otherwise.

The results are shown in Table 12. We can see that the coefficient in front of the interaction term is negatively estimated with statistical significance. This suggests that, for banks with longer period until the timestamp of the deposit flows, the response to deposit discussion is stronger. This corroborates our argument that our deposit discussion measures capture information released from earnings calls, distinct from other channels of news. This finding supports our overall empirical approach that novel insights can be captured using textual analysis on earnings calls.

4.4.2. Connection with the Literature on Bank Illiquidity

In the following analysis, we examine how the depositors' response depends on the fraction of uninsured deposits. This line of testing extends the work by Chen et al. (2024) who highlight that larger liquidity transformation is associated with higher fragility. We follow one of their measures of liquidity mismatch, that is the degree of uninsured deposits out of total deposit financing, and we examine how this relates to our main finding on depositor's responsiveness to earnings call discussions.

We start by simply controlling for the uninsured deposit share and repeating our main regressions. The first four columns of Table 13 show that the fraction of uninsured deposits predicts deposit outflows with t-statistics larger than -6.5 for all specifications. This indicates that when uninsured deposits are at high levels, they tend to fall subsequently. The estimated coefficients in front of our deposit discussion scores show that the predictability pattern largely remains, while only the *depositmatch* coefficient becomes statistically weak. This means that the effect of deposit discussions is not explained away by the high fraction of uninsured deposit funding. In other words, bank management discusses deposit topics for reasons other than just high levels of liquidity mismatch.

Next, we run the regressions with the interaction between our deposit discussion score and uninsured deposit share. The results are shown in the fifth through eighth columns. We can see that the coefficients in front of the interacted terms are negatively estimated. Interestingly, the coefficients in front of the deposit discussion metrics are positively estimated. Hence, we compute the plus and minus one standard deviation around the mean for both key variables to understand the directional impact. $uninsuredfraction$ has mean 0.49

and standard deviation of 0.19, while the deposit discussion metric has mean zero and unit standard deviation. Given a high value of $uninsuredfraction$, namely at the median value plus one standard deviation, the impact of increasing $depositcnt$ from zero to one results in a decrease in uninsured deposits by 0.005. Given the median value of $uninsuredfraction$, the marginal effect of deposit talks is a decrease by 0.001, and finally given a small value $uninsuredfraction$ equal to the median minus one standard deviation, the marginal effect flips sign: an increase by 0.001.

This means that when a bank's uninsured deposit funding is high, the deposit flow response to deposit discussions is negative and economically large. When a bank's uninsured deposit funding is low, the response can flip, indicating a deposit inflow. However, in terms of the economic magnitude, the interesting result is that the deposit outflow response becomes significant when uninsured deposit share is high. Hence, the takeaway is that depositors' responsiveness to earnings calls is especially high when the bank is potentially fragile because of high liquidity mismatch.

We then follow up the previous analysis by specifically looking at the size of average balance of the uninsured deposits instead of calculating the overall uninsured deposit funding of a bank. In other words, we divide the uninsured deposit amount by the number of accounts that are uninsured. The intention is to capture how much uninsured money is at risk for the affected account holders.

The results in Table 14 indicate that the size of average uninsured balances negatively predicts uninsured deposit flows. At the same time, we find that our deposit discussion metrics are mostly robust and unaffected across most of the specifications. Hence, our deposit discussion channel is quite distinct from the size of uninsured account balances. The results from Table 13 indicate, instead, that having many uninsured account holders can substantially affect the depositor's responsiveness.

4.4.3. Do Banks Pro-actively Discuss Deposit Topics?

Lastly, we conduct analysis to test whether or not there are any notable differences between the discussions in the presentation section of the transcript or those in the question-and-answer (Q&A) section. It is plausible that the former section of the call is more driven by what bank management strategically discloses to stakeholders, while the latter section of the call is more driven by what stakeholders actually hope to listen from management. It is possible that there is a difference between the two.

However, we do not find strong evidence of any notable difference at least in terms of response to deposit discussions. We dissect the call transcript to the presentation part and the Q&A part. We then re-construct the text-based discussion metrics based on the text restricted each part. The first four columns of Table 15 show the results with the presentation-based metrics, while the latter four columns are based on the Q&A-based metrics. We can see that the deposit discussion metrics seem to load negatively for both sections, but the statistical power is absent for the Q&A-based metrics. Bank's management seems to voluntarily provide updates on deposit topics, which would be followed by depositor's response. The slightly weak statistical power of the Q&A section indicates that depositors' responsiveness is not as strong based on the dialogue from this two-way conversation.

Without observing a meaningful difference between the two, we would conclude that banks do not seem to be trying to omit deposit-related concerns from the presentation part of the call, which happens before analysts participating in the calls raise questions. This finding points to the direction that there is valuable news that can be learned from earnings calls because bank management is already active in disclosing information about deposits.

5. Conclusion

This paper provides evidence that banks' management communicates informative signals about depositor behavior during earnings call meetings. We find that more substantial deposit-related discussions are followed by outflows of uninsured deposit balances, while this reaction is not observed for insured deposit balances. We also find that uninsured deposit flow response is stronger when negative news arrive, indicating that the flow response is related to depositors' concerns. Further examination of the cross-sectional heterogeneity reveals that the deposit flow response is stronger for banks more serving urban population, higher expenditure population, and areas with higher focus in the information sector. We also find evidence that when deposit discussions rise, banks make up for the lost deposit funding by raising rates to obtain funding from sophisticated investors.

These pieces of evidence point to the direction that there is valuable information that is revealed by banks' management about what depositors think about their depositing institutions. This information provides signals to uninsured deposit account holders who would be sensitive to the health of their depositing institutions. This finding highlights that depositors, especially those that are active and responsive, pay attention to news from this type of communication channel.

Our study also sheds light on future potential research questions. For instance, do banks' management strategically choose to reveal more information about specific topics than others? Are signals of certain topics more noisy or misleading than other topics? Given the potential instability of deposit funding, our study contributes to the burgeoning literature on the use of textual analysis to understand the interaction between banks' management and their customers.

References

Amoozegar, A., Berger, D., Cao, X., Pukthuanthong, K., 2020. Earnings conference calls and institutional monitoring: Evidence from textual analysis. Journal of Financial Research 43, 5–36.

Benmelech, E., Yang, J., Zator, M., 2023. Bank branch density and bank runs. NBER Working Paper Series No. 31462.

Blei, D.M., Ng, A.Y., Jordan, M.I., 2003. Latent dirichlet allocation. The Journal of Machine Learning Research 3, 993–1022.

Blickle, K., Li, J., Lu, X., Ma, Y., 2024. The dynamics of deposit flightiness and its impact on financial stability. Working Paper .

Bochkay, K., Hales, J., Chava, S., 2020. Hyperbole or reality? investor response to extreme language in earnings conference calls. The Accounting Review 95, 31–60.

Bolton, P., Li, Y., Wang, N., Yang, J., 2025. Dynamic banking and the value of deposits. The Journal of Finance 80, 2063–2105.

Bushman, R.M., Hendricks, B.E., Williams, C.D., 2016. Bank competition: Measurement, decision-making, and risk-taking. Journal of Accounting Research 54, 777–826.

Call, A.C., Flam, R.W., Lee, J.A., Sharp, N.Y., 2024. Managers' use of humor on public earnings conference calls. Review of Accounting Studies 29, 2650–2687.

Chen, J.V., Nagar, V., Schoenfeld, J., 2018. Manager-analyst conversations in earnings conference calls. Review of Accounting Studies 23, 1315–1354.

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

Chen, Q., Goldstein, I., Huang, Z., Vashishtha, R., 2024. Liquidity transformation and fragility in the u.s. banking sector. The Journal of Finance 79, 3985–4036.

Cipriani, M., Eisenbach, T.M., Kovner, A., 2024. Tracing bank runs in real time. Federal Reserve Bank of New York Staff Reports No. 1104.

Cook, T.R., Hansen, A.L., Kazinnik, S., McAdam, P., 2023. Evaluating local language models: An application to financial earnings calls. Research Working Papers RWP 23-12.

Cookson, J.A., Fox, C., Gil-Bazo, J., Imbet, J.F., Schiller, C., 2024. Social media as a bank run catalyst. Working Paper .

De Amicis, C., Falconieri, S., Tastan, M., 2021. Sentiment analysis and gender differences in earnings conference calls. Journal of Corporate Finance 71, 101809.

DeMarzo, P.M., Krishnamurthy, A., Nagel, S., 2024. Interest rate risk in banking. NBER Working Paper Series No. 33308.

Diamond, D.W., Dybvig, P.H., 1983. Bank runs, deposit insurance, and liquidity. Journal of Political Economy 91, 401–419.

Drechsler, I., Savov, A., Schnabl, P., 2017. The deposits channel of monetary policy. The Quarterly Journal of Economics 132, 1819–1876.

Drechsler, I., Savov, A., Schnabl, P., 2021. Banking on deposits: Maturity transformation without interest rate risk. The Journal of Finance 76, 1091–1143.

Egan, M., Hortaçsu, A., Matvos, G., 2017. Deposit competition and financial fragility: Evidence from the us banking sector. American Economic Review 107, 169–216.

Emin, M., James, C., Li, T., 2025. Variable deposit betas and bank exposure to interest rate risk. Journal of Financial Intermediation, *forthcoming* .

Erel, I., Liebersohn, J., Yannelis, C., Earnest, S., 2023. Monetary policy transmission through online banks. NBER Working Paper Series No. 31380.

Goldstein, I., Pauzner, A., 2005. Demand–deposit contracts and the probability of bank runs. The Journal of Finance 60, 1293–1327.

Iyer, R., Puri, M., 2012. Understanding bank runs: The importance of depositor-bank relationships and networks. American Economic Review 102, 1414–45.

Jiang, E.X., Matvos, G., Piskorski, T., Seru, A., 2024. Monetary tightening and u.s. bank fragility in 2023: Mark-to-market losses and uninsured depositor runs? Journal of Financial Economics 159, 103899.

Jiang, L., Levine, R., Lin, C., Xie, W., 2022. Deposit supply and bank transparency. Management Science 68, 3834–3855.

Koont, N., Santos, T., Zingales, L., 2024. Destabilizing digital "bank walks". NBER Working Paper Series No. 32601.

Lee, J., 2016. Can investors detect managers' lack of spontaneity? adherence to predetermined scripts during earnings conference calls. The Accounting Review 91, 229–250.

Martin, C., Puri, M., Ufier, A., 2022. Deposit inflows and outflows in failing banks: The role of deposit insurance. FDIC Working Paper Series .

Price, S.M., Doran, J.S., Peterson, D.R., Bliss, B.A., 2012. Earnings conference calls and stock returns: The incremental informativeness of textual tone. Journal of Banking & Finance 36, 992–1011.

Rennekamp, K.M., Sethuraman, M., Steenhoven, B.A., 2022. Engagement in earnings conference calls. Journal of Accounting and Economics 74, 101498.

Zhang, S., 2023. Analyzing bank cybersecurity disclosures using machine learning. Working paper .

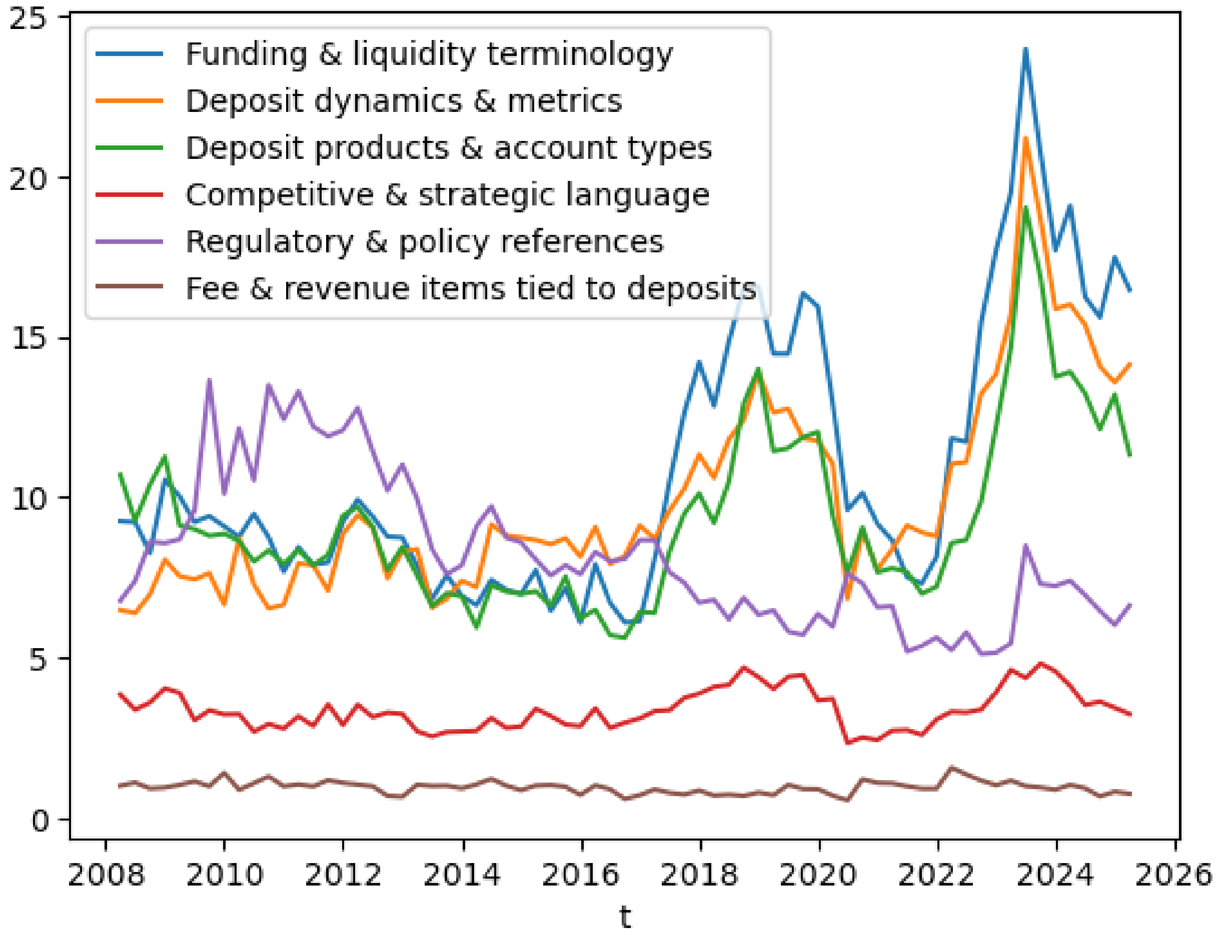

Figure 1: Across-bank average counts of deposit-related keywords by sub-category. Notes: For each transcript, the number of occurrences that keywords associated with each category is recorded. Shown in the figure are the averages across banks for a given quarter.

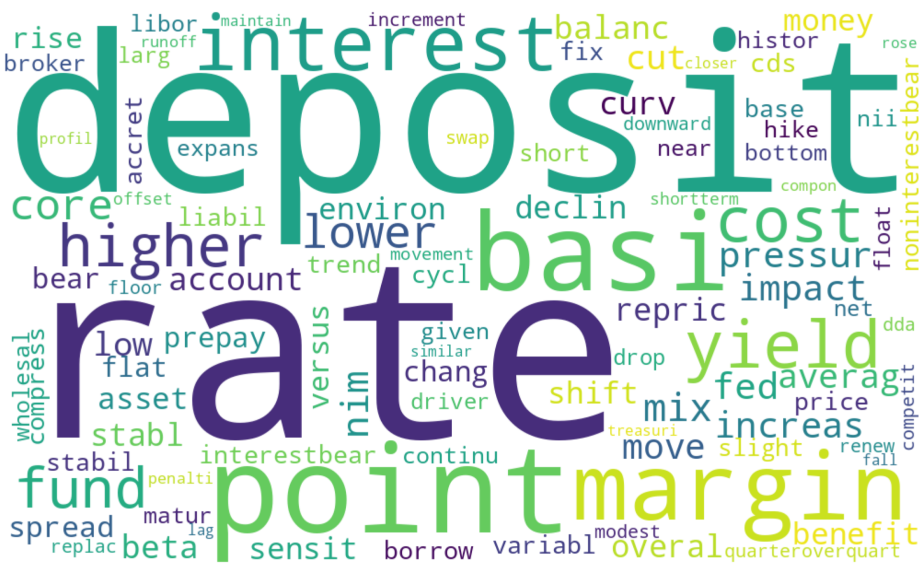

Figure 2: Word cloud of the deposit topic from LDA. *Notes*: Running LDA on 500 random draws of transcripts gives 20 topics, from which the deposit-related topic is selected based on whether the word "deposit" appears as the top-probability word.

Key strands surfaced by the 100 highest-weight terms

Cluster of high-probability terms	What they reveal
rate · point · basis · higher · lower · Fed · curve · shift · flat · rise · decline	A macro frame: policy-rate hikes/cuts and yield-curve shape changes drive the discussion.
deposit · core · fund · mix · money · liability · broker/wholesale hints	Focus on the *composition* of funding — retaining sticky core deposits vs. costlier alternatives.
cost · margin · NIM · spread · compress · pressure · impact · benefit	Profit lens: how funding-cost moves squeeze or enlarge NIM and spread.
beta · repric(e) · sensit(ivity) · lag · offset · movement	Mechanics of *deposit-rate pass-through* — speed/extent (deposit β) and timing lags.
yield · asset · swap · CDS · variable · fixed · maturity · prepay · runoff	Balance-sheet tactics: asset-yield lift, hedging with swaps/CDS, managing term structure, prepayments, and runoff profiles.
stable · trend · environment · overall · cycle	Narrative about stability vs. volatility and quarter-over-quarter trends.

Figure 3: ChatGPT's interpretation of the deposit topic generated from LDA. *Notes*: LDA produces as output a list of words and associated probabilities. This output is fed into ChatGPT for interpretation. The left column shows ChatGPT's grouping of words, and thr right column shows the potential meaning of each group.

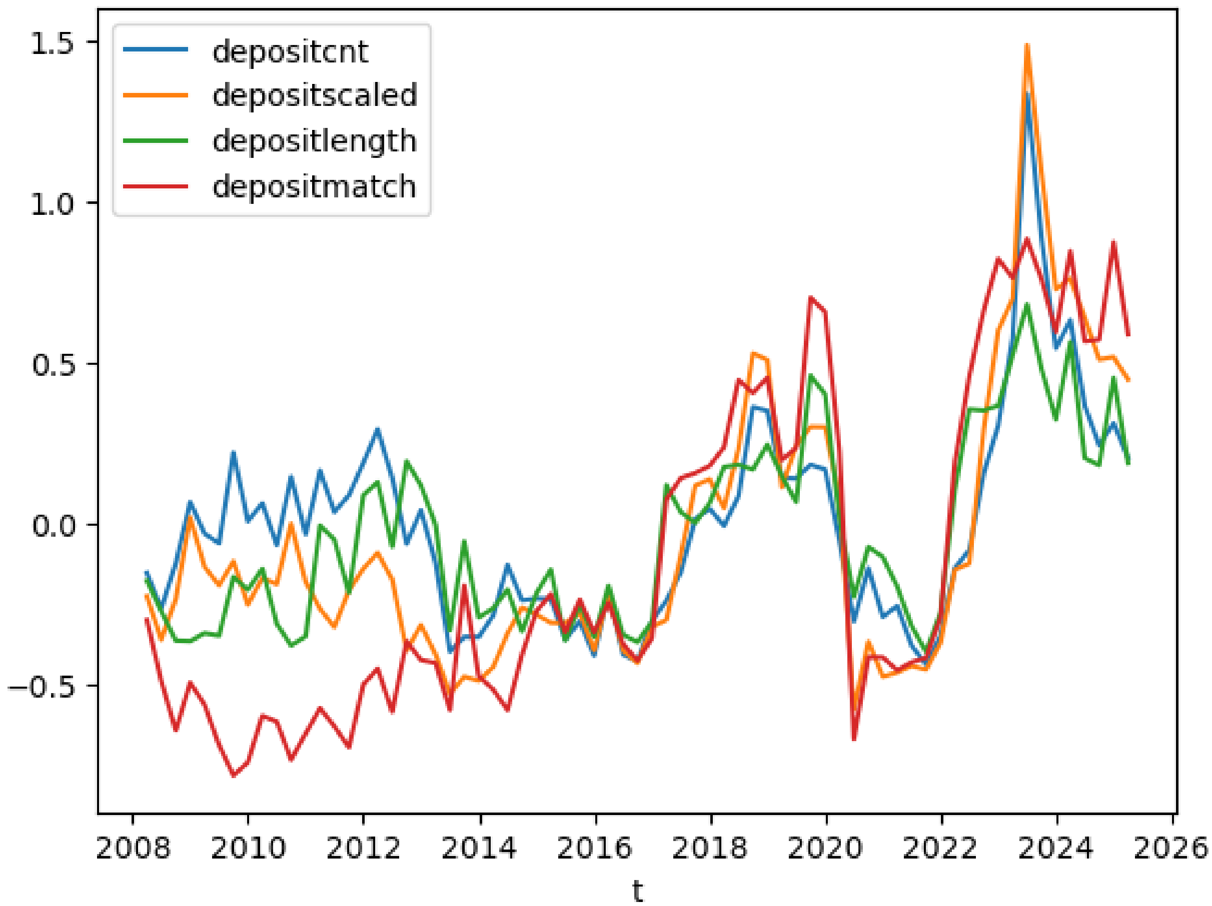

Figure 4: Across-bank average of deposit discussion metrics. *Notes*: *depositcnt* and *depositscaled* are the metrics based on keyword search, and *depositlength* and *depositmatch* are based on the deposit topic generated by LDA. Shown on the graph are the averages across banks for each quarter.

Variable	Obs	Mean	Std. dev.	25th%	75th%
deposits	8614	71,700,000	243,000,000	4,119,267	30,000,000
uninsured deposits	8552	43,800,000	160,000,000	1,719,799	14,700,000
roa	8609	0.0026	0.0036	0.002	0.0034
log assets	8614	16.57	1.66	15.43	17.49
net interest margin	8609	0.008	0.0033	0.0068	0.0087
loan-loss provision ratio	8553	0.0012	0.0031	0.0001	0.0011
ppop to loans	8553	0.008	0.0551	0.0049	0.0078
debt to equity	8614	8.593	13.3561	6.8224	9.5112
loan to deposit ratio	8614	0.847	0.2363	0.7591	0.9762
core deposit to asset	8614	0.599	0.15	0.512	0.706
non-performing loans	8170	953,520	4,453,419	18,033	203,874
depositylield	8609	0.0021	0.0022	0.0005	0.0029

Table 1: Summary Statistics. *Notes*: the sample period used in this study covers from the first quarter of 2008 to the fourth quarter of 2024. Deposit discussion metrics are standardized to have zero mean and unit standard deviation and hence are omitted in this table.

Category: Competitive & strategic language	Category: Deposit dynamics & metrics	Category: Deposit products & account types
bundle checking	funding mix	cd
bundle saving	share of wallet	checking account
customer acqui	deposit grow	current account
digital account	deposit flow	dda
introductory rate	deposit migrat	mmda
online account	deposit runoff	money market account
promo rate	deposit attrition	now account
rate cap	deposit retention	passbook saving
rate control	deposit stick	saving account
teaser rate	deposit stab	sweep account
deposit compet	average deposit	demand deposit
deposit pressure pric	ending deposit	certificate of deposit
deposit campaign	period deposit	time deposit
deposit acqui	deposit balance	term deposit
deposit gather	deposit mix	fixed deposit
mobile deposit	deposit compos	interest bearing
cross sell deposit	deposit share	ib deposit
	deposit season	escrow deposit
	deposit forecast	trust deposit
	deposit outlook	fiduciary deposit
	deposit franchise	municipal deposit
	relation deposit	public fund deposit
		commercial deposit
		corporate deposit
		treasury deposit
		consumer deposit
		retail deposit
		small business deposit
		wholesale deposit
		online deposit
		digital deposit
		direct deposit

Category: Funding & liquidity terminology	Category: Regulatory & policy references	Category: Fee & revenue items tied to deposits
core fund	fdic assessment	nsf fee
cost fund	ics	overdraft fee
fdic insur	insured cash sweep	deposit service charge
funding base	reg d	deposit fee
nsfr	reserve requirement	
stable funding	brokered deposit	
core deposit	reciprocal deposit	
deposit cost	deposit guarantee	
deposit beta		
deposit rate		
deposit pric		
deposit repric		
deposit elasticity		
deposit yield		
interest expense deposit		
deposit duration		
deposit sensitivity		
deposit insur		
collateralized deposit		
collateralised deposit		
pledged deposit		

Table 2: List of deposit keywords constructed by asking ChatGPT. Note that when constructing the deposit discussion metrics, we simply count across keywords without consideration of the category. We also replace all the keywords containing the word 'deposit' simply by the keyword 'deposit'.

	Panel A Dependent: uninsured deposit change next period				Panel B Dependent: contemporaneous uninsured deposit change				Panel C Dependent: uninsured deposit change two-period ahead			
depositcnt	−0.0018***				−0.0026***				−0.0011			
	(−2.785)				(−2.853)				(−1.537)			
depositscaled		−0.0010				−0.0010				−0.0003		
		(−1.497)				(−1.208)				(−0.373)		
depositlength			−0.0026***				−0.0016				−0.0032***	
			(−3.171)				(−1.381)				(−3.681)	
depositmatch				−0.0019**				−0.0000				−0.0023**
				(−2.323)				(−0.008)				(−2.307)
roa	−0.0212	−0.0219	−0.0225	−0.0214	0.2707	0.2685	0.2660	0.2654	0.1474	0.1458	0.1488	0.1498
	(−0.140)	(−0.144)	(−0.145)	(−0.138)	(0.825)	(0.821)	(0.806)	(0.810)	(0.804)	(0.793)	(0.795)	(0.802)
logassets	−0.0224***	−0.0228***	−0.0223***	−0.0230***	0.0147***	0.0142***	0.0145***	0.0142***	−0.0190***	−0.0193***	−0.0187***	−0.0195***
	(−6.076)	(−6.068)	(−6.109)	(−6.116)	(3.698)	(3.587)	(3.631)	(3.599)	(−4.839)	(−4.845)	(−4.837)	(−4.905)
nim	0.4669	0.4757	0.4985	0.4788	2.5631	2.5998	2.6287	2.6398	0.8977**	0.9170**	0.9087**	0.8845**
	(0.985)	(0.997)	(1.051)	(1.006)	(1.365)	(1.378)	(1.392)	(1.399)	(2.066)	(2.086)	(2.090)	(2.020)
loanlossprovisionratio	−0.5500**	−0.5580**	−0.5634**	−0.5705**	0.8256	0.8117	0.7999	0.8067	−0.1149	−0.1215	−0.1238	−0.1319
	(−2.182)	(−2.208)	(−2.242)	(−2.261)	(1.407)	(1.380)	(1.349)	(1.366)	(−0.364)	(−0.386)	(−0.390)	(−0.418)
ppop_loans	−0.0160***	−0.0157***	−0.0155***	−0.0155***	0.0481***	0.0485***	0.0487***	0.0486***	0.0308***	0.0310***	0.0312***	0.0311***
	(−3.606)	(−3.542)	(−3.537)	(−3.508)	(4.840)	(4.886)	(4.900)	(4.897)	(6.875)	(6.896)	(7.089)	(6.981)
debttoequity	−0.0004	−0.0004	−0.0004	−0.0004	0.0016**	0.0016**	0.0016**	0.0016**	−0.0002	−0.0002	−0.0002	−0.0002
	(−0.873)	(−0.876)	(−0.862)	(−0.868)	(2.391)	(2.384)	(2.395)	(2.384)	(−0.501)	(−0.506)	(−0.479)	(−0.484)
loantodepositratio	0.0236**	0.0235**	0.0237**	0.0240**	−0.0639***	−0.0642***	−0.0642***	−0.0645***	0.0090	0.0089	0.0095	0.0098
	(2.110)	(2.095)	(2.144)	(2.144)	(−5.473)	(−5.478)	(−5.467)	(−5.533)	(0.761)	(0.744)	(0.811)	(0.822)
coredeposits_assets	0.0103	0.0106	0.0100	0.0104	−0.0681***	−0.0677***	−0.0681***	−0.0678***	0.0103	0.0104	0.0098	0.0103
	(0.973)	(0.998)	(0.956)	(0.985)	(−6.726)	(−6.627)	(−6.717)	(−6.608)	(1.107)	(1.114)	(1.083)	(1.124)
dln_deposits_4q	0.0172	0.0177	0.0174	0.0178					0.0084	0.0087	0.0083	0.0086
	(0.834)	(0.850)	(0.846)	(0.857)					(0.481)	(0.494)	(0.481)	(0.488)
dln_deposits_4q_m1					−0.0284**	−0.0281**	−0.0280**	−0.0280**				
					(−2.009)	(−2.014)	(−2.003)	(−2.024)				
Constant	0.3576***	0.3632***	0.3555***	0.3661***	−0.1715**	−0.1636**	−0.1680**	−0.1636**	0.3074***	0.3107***	0.3010***	0.3140***
	(5.570)	(5.583)	(5.579)	(5.626)	(−2.501)	(−2.405)	(−2.445)	(−2.415)	(4.740)	(4.748)	(4.717)	(4.801)
Observations	8,276	8,276	8,276	8,276	8,442	8,442	8,442	8,442	8,080	8,080	8,080	8,080
Adjusted R-squared	0.117	0.117	0.118	0.117	0.128	0.127	0.128	0.127	0.137	0.136	0.138	0.137

Table 3: Regression of uninsured deposit flows on deposit discussion metrics. *Notes*: *depositcnt* and *depositscaled* are the metrics based on keyword search, and *depositlength* and *depositmatch* are based on the deposit topic generated by LDA. In Panel A, the dependent variable is the change from t to t+1. In Panel B, the dependent variable is the change from t-1 to t, while in Panel C the change is from t+1 to t+2. Fixed effects include bank and year- quarter. Standard errors are clustered at the bank-level.

	Panel A Dependent: insured deposit share change next period				Panel B Dependent: total deposit share change next period			
depositcnt	0.0010				−0.0008			
	(1.392)				(−0.817)			
depositscaled		0.0015**				0.0004		
		(2.187)				(0.423)		
depositlength			0.0000				−0.0026**	
			(0.045)				(−2.169)	
depositmatch				0.0005				−0.0013
				(0.699)				(−1.050)
roa	−0.2016	−0.2045	−0.1995	−0.2007	−0.2238	−0.2271	−0.2224	−0.2223
	(−0.565)	(−0.575)	(−0.559)	(−0.562)	(−0.619)	(−0.628)	(−0.618)	(−0.617)
logassets	−0.0238***	−0.0236***	−0.0236***	−0.0235***	−0.0450***	−0.0451***	−0.0447***	−0.0453***
	(−6.846)	(−6.877)	(−6.847)	(−6.841)	(−7.338)	(−7.348)	(−7.355)	(−7.345)
nim	1.1025	1.1322*	1.0740	1.0848	1.6009	1.6412	1.5986	1.5961
	(1.632)	(1.675)	(1.584)	(1.602)	(1.570)	(1.608)	(1.567)	(1.564)
loanlossprovisionratio	−2.1703	−2.1701	−2.1624	−2.1598	−2.7175	−2.7261	−2.7234	−2.7302
	(−1.284)	(−1.285)	(−1.280)	(−1.277)	(−1.500)	(−1.505)	(−1.510)	(−1.511)
ppop_loans	−0.0169***	−0.0169***	−0.0171***	−0.0171***	−0.0321***	−0.0319***	−0.0319***	−0.0319***
	(−3.390)	(−3.396)	(−3.430)	(−3.440)	(−4.227)	(−4.196)	(−4.212)	(−4.199)
debttoequity	−0.0003	−0.0003	−0.0003	−0.0003	−0.0005	−0.0005	−0.0005	−0.0005
	(−0.895)	(−0.903)	(−0.890)	(−0.899)	(−0.995)	(−0.993)	(−0.985)	(−0.986)
loantodepositratio	0.0360***	0.0357***	0.0363***	0.0361***	0.0596***	0.0592***	0.0600***	0.0600***
	(5.098)	(5.094)	(5.077)	(5.049)	(4.354)	(4.331)	(4.403)	(4.352)
coredeposits_assets	0.0094	0.0092	0.0093	0.0094	0.0216	0.0216	0.0212	0.0216
	(1.046)	(1.027)	(1.028)	(1.031)	(1.345)	(1.353)	(1.332)	(1.350)
dln_deposits_4q	0.0689**	0.0690**	0.0687**	0.0687**	0.0775***	0.0778***	0.0776***	0.0777***
	(2.299)	(2.299)	(2.282)	(2.282)	(3.404)	(3.433)	(3.401)	(3.424)
Constant	0.3631***	0.3603***	0.3600***	0.3592***	0.6980***	0.7007***	0.6933***	0.7024***
	(6.211)	(6.224)	(6.207)	(6.191)	(6.722)	(6.733)	(6.720)	(6.733)
Observations	8,342	8,342	8,342	8,342	8,342	8,342	8,342	8,342
Adjusted R-squared	0.096	0.096	0.095	0.095	0.085	0.085	0.086	0.085

Table 4: Regressions with insured deposit change and total deposit change. Notes: In Column A the dependent variable is insured deposit changes from t to t+1, while in column B it is total deposit changes from t to t+1.

	Panel A Dependent: uninsured deposit change, next period				Panel B Dependent: uninsured deposit change, contemporaneous			
depositcntXiave_negsentiment	-0.0029^{***}				-0.0021^{*}			
	(-2.968)				(-1.864)			
depositscaledXiave_negsentiment		-0.0030^{***}				-0.0002		
		(-2.661)				(-0.179)		
depositlengthXiave_negsentiment			-0.0024^{**}				-0.0020^{*}	
			(-2.363)				(-1.905)	
depositmatchXiave_negsentiment				-0.0025^{**}				-0.0001
				(-2.014)				(-0.106)
depositcnt	-0.0004				-0.0016			
	(-0.465)				(-1.601)			
depositscaled		0.0005				-0.0009		
		(0.559)				(-0.915)		
depositlength			-0.0014				-0.0003	
			(-1.354)				(-0.243)	
depositmatch				-0.0005				0.0002
				(-0.495)				(0.202)
iave_negsentiment	0.0004	0.0004	0.0006	0.0005	-0.0027^{**}	-0.0027^{**}	-0.0026^{**}	-0.0027^{**}
	(0.366)	(0.356)	(0.524)	(0.475)	(-2.417)	(-2.387)	(-2.361)	(-2.410)
roa	-0.0174	-0.0169	-0.0195	-0.0130	0.2657	0.2611	0.2607	0.2579
	(-0.114)	(-0.111)	(-0.125)	(-0.082)	(0.814)	(0.799)	(0.795)	(0.787)
logassets	-0.0223^{***}	-0.0227^{***}	-0.0222^{***}	-0.0230^{***}	0.0150^{***}	0.0144^{***}	0.0146^{***}	0.0144^{***}
	(-6.038)	(-6.030)	(-6.073)	(-6.085)	(3.751)	(3.634)	(3.665)	(3.655)
nim	0.4417	0.4526	0.4776	0.4584	2.5321	2.5713	2.6022	2.6148
	(0.944)	(0.956)	(1.024)	(0.975)	(1.351)	(1.364)	(1.378)	(1.386)
loanlossprovisionratio	-0.5840^{**}	-0.5924^{**}	-0.6071^{**}	-0.6143^{**}	0.8772	0.8690	0.8465	0.8658
	(-2.340)	(-2.369)	(-2.455)	(-2.475)	(1.466)	(1.446)	(1.400)	(1.434)
$ppop_{loans}$	-0.0159^{***}	-0.0160^{***}	-0.0154^{***}	-0.0157^{***}	0.0484^{***}	0.0486^{***}	0.0490^{***}	0.0488^{***}
	(-3.606)	(-3.617)	(-3.532)	(-3.572)	(4.875)	(4.917)	(4.935)	(4.928)
debttoequity	-0.0003	-0.0003	-0.0003	-0.0003	0.0016^{**}	0.0016^{**}	0.0016^{**}	0.0016^{**}
	(-0.826)	(-0.839)	(-0.793)	(-0.803)	(2.359)	(2.352)	(2.366)	(2.352)
loantodepositratio	0.0240^{**}	0.0238^{**}	0.0243^{**}	0.0245^{**}	-0.0642^{***}	-0.0645^{***}	-0.0643^{***}	-0.0649^{***}
	(2.175)	(2.131)	(2.262)	(2.225)	(-5.500)	(-5.495)	(-5.457)	(-5.566)
coredeposits_assets	0.0108	0.0112	0.0104	0.0109	-0.0681^{***}	-0.0678^{***}	-0.0683^{***}	-0.0679^{***}
	(1.024)	(1.064)	(1.003)	(1.046)	(-6.761)	(-6.666)	(-6.734)	(-6.647)
dln_deposits_4q	0.0162	0.0171	0.0170	0.0177				
	(0.794)	(0.825)	(0.828)	(0.852)				
dln_deposits_4q_m1					-0.0295^{**}	-0.0285^{**}	-0.0287^{**}	-0.0284^{**}
					(-2.063)	(-2.036)	(-2.042)	(-2.048)
Constant	0.3548^{***}	0.3616^{***}	0.3533^{***}	0.3653^{***}	-0.1739^{**}	-0.1650^{**}	-0.1690^{**}	-0.1651^{**}
	(5.556)	(5.567)	(5.569)	(5.618)	(-2.524)	(-2.422)	(-2.453)	(-2.438)
Observations	8,273	8,273	8,273	8,273	8,439	8,439	8,439	8,439
Adjusted R-squared	0.118	0.118	0.119	0.118	0.129	0.128	0.128	0.128

Table 5: Interaction with sentiment. *Notes:* The independent variable of interest is the interaction between the deposit discussion metric and an indicator variable on whether the sentiment of the given transcript is below the median for the corresponding quarter. In Panel A, the deposit flow is measured from t to t+1, while in Panel B the flow is measured from t-1 to t.

	Panel A				Panel B			
	Dependent: uninsured deposit change, next period				Dependent: uninsured deposit change, contemporaneous			
depositcntXminusroa	−0.2644				−0.7774*			
	(−1.326)				(−1.924)			
depositscaledXminusroa		−0.3245*				−0.8756**		
		(−1.653)				(−2.202)		
depositlengthXminusroa			0.0091				−0.7156	
			(0.047)				(−1.274)	
depositmatchXminusroa				0.1098				−0.3544
				(0.997)				(−0.791)
depositcnt	−0.0024***				−0.0045***			
	(−2.794)				(−3.443)			
depositscaled		−0.0018**				−0.0031**		
		(−2.053)				(−2.497)		
depositlength			−0.0026**				−0.0035*	
			(−2.391)				(−1.926)	
depositmatch				−0.0016*				−0.0009
				(−1.805)				(−0.636)
minusroa	−0.0217	−0.0531	0.0263	0.0784	−0.4004	−0.4683	−0.5537	−0.4409
	(−0.148)	(−0.354)	(0.141)	(0.427)	(−1.290)	(−1.377)	(−1.280)	(−0.943)
logassets	−0.0224***	−0.0227***	−0.0223***	−0.0230***	0.0148***	0.0144***	0.0146***	0.0143***
	(−6.022)	(−6.015)	(−6.108)	(−6.113)	(3.737)	(3.665)	(3.662)	(3.618)
nim	0.4719	0.4447	0.4988	0.4951	2.5621	2.4891	2.6179	2.5891
	(1.001)	(0.932)	(1.052)	(1.046)	(1.367)	(1.310)	(1.391)	(1.359)
loanlossprovisionratio	−0.5333**	−0.5551**	−0.5639**	−0.5589**	0.9153	0.8775	0.8141	0.7572
	(−2.056)	(−2.170)	(−2.229)	(−2.186)	(1.550)	(1.525)	(1.403)	(1.388)
ppop$_{l}oans$	−0.0158***	−0.0152***	−0.0155***	−0.0158***	0.0485***	0.0502***	0.0493***	0.0496***
	(−3.351)	(−3.004)	(−3.527)	(−3.687)	(5.474)	(6.331)	(5.442)	(5.598)
debttoequity	−0.0004	−0.0003	−0.0004	−0.0004	0.0016**	0.0017**	0.0016**	0.0016**
	(−0.861)	(−0.829)	(−0.862)	(−0.882)	(2.411)	(2.442)	(2.405)	(2.401)
loantodepositratio	0.0234**	0.0235**	0.0237**	0.0240**	−0.0644***	−0.0640***	−0.0649***	−0.0646***
	(2.071)	(2.077)	(2.147)	(2.154)	(−5.618)	(−5.520)	(−5.575)	(−5.554)
coredeposits$_{a}ssets$	0.0106	0.0110	0.0100	0.0103	−0.0669***	−0.0662***	−0.0675***	−0.0675***
	(1.006)	(1.042)	(0.954)	(0.976)	(−6.653)	(−6.598)	(−6.571)	(−6.565)
dln_deposits_4q	0.0167	0.0169	0.0174	0.0180				
	(0.811)	(0.815)	(0.846)	(0.866)				
dln_deposits_4q_m1					−0.0293**	−0.0302**	−0.0283**	−0.0285**
					(−2.043)	(−2.140)	(−2.008)	(−2.048)
Constant	0.3569***	0.3618***	0.3556***	0.3664***	−0.1740**	−0.1682**	−0.1706**	−0.1648**
	(5.528)	(5.538)	(5.577)	(5.620)	(−2.551)	(−2.497)	(−2.482)	(−2.436)
Observations	8,276	8,276	8,276	8,276	8,442	8,442	8,442	8,442
Adjusted R-squared	0.117	0.117	0.118	0.117	0.131	0.130	0.129	0.128

Table 6: Interaction with the return on assets. *Notes:* The independent variable of interest is the interaction between the deposit discussion metric and the return on assets. In Panel A, the deposit flow is measured from t to t+1, while in Panel B the flow is measured from t-1 to t.

	Panel A				Panel B			
	Dependent: chg. log(NPL) next period				Dependent: chg. log(NPL) two-period-ahead			
depositcnt	0.0122*				0.0151**			
	(1.950)				(2.472)			
depositscaled		0.0098				0.0101*		
		(1.278)				(1.708)		
depositlength			0.0102				0.0168**	
			(1.605)				(2.212)	
depositmatch				0.0036				0.0125**
				(0.443)				(2.022)
roa	2.2107	2.2301	2.2593	2.2653	−0.5158	−0.4894	−0.4626	−0.4534
	(1.178)	(1.182)	(1.190)	(1.193)	(−0.300)	(−0.281)	(−0.267)	(−0.259)
logassets	−0.0130	−0.0106	−0.0126	−0.0104	−0.0332*	−0.0301*	−0.0335*	−0.0290
	(−0.695)	(−0.573)	(−0.672)	(−0.560)	(−1.837)	(−1.695)	(−1.842)	(−1.623)
nim	−1.0031	−0.9315	−1.2828	−1.2680	−3.3665	−3.3629	−3.7217	−3.5557
	(−0.377)	(−0.354)	(−0.474)	(−0.467)	(−1.044)	(−1.037)	(−1.152)	(−1.089)
loanlossprovisionratio	−2.4159	−2.3463	−2.2195	−2.2176	−2.3841	−2.2752	−2.1359	−2.0875
	(−0.898)	(−0.868)	(−0.809)	(−0.810)	(−0.906)	(−0.854)	(−0.797)	(−0.775)
ppop_loans	−0.5108	−0.5231	−0.5347	−0.5330	−0.3999	−0.4154	−0.4307	−0.4302
	(−0.953)	(−0.961)	(−0.991)	(−0.982)	(−0.536)	(−0.551)	(−0.581)	(−0.572)
debttoequity	0.0040	0.0040	0.0040	0.0040	0.0013	0.0013	0.0012	0.0012
	(1.429)	(1.437)	(1.418)	(1.442)	(0.525)	(0.529)	(0.500)	(0.496)
loantodepositratio	0.0610	0.0596	0.0614	0.0621	0.0117	0.0111	0.0109	0.0083
	(1.636)	(1.602)	(1.637)	(1.642)	(0.271)	(0.261)	(0.252)	(0.197)
coredeposits_assets	0.1183*	0.1160*	0.1187**	0.1162*	0.0984	0.0956	0.0999	0.0969
	(1.969)	(1.929)	(1.974)	(1.927)	(1.646)	(1.593)	(1.642)	(1.602)
dln_deposits_4q	0.1022	0.1013	0.0993	0.0990	−0.0227	−0.0246	−0.0255	−0.0260
	(0.981)	(0.983)	(0.958)	(0.967)	(−0.177)	(−0.194)	(−0.199)	(−0.205)
Constant	0.0807	0.0421	0.0761	0.0388	0.5233	0.4742	0.5315	0.4592
	(0.245)	(0.129)	(0.230)	(0.119)	(1.613)	(1.480)	(1.635)	(1.429)
Observations	7,833	7,833	7,833	7,833	7,660	7,660	7,660	7,660
Adjusted R-squared	0.037	0.037	0.037	0.037	0.040	0.040	0.040	0.040

Table 7: Prediction of non-performing loan growth. *Notes:* The dependent variable is the change in the log of the non-performing loan balance. In Panel A the change in log is measured from t to t+1, while in Panel B the change is measured from t+1 to t+2.

	Panel A				Panel B				Panel C			
depositcntXurban	−0.0034*											
	(−1.878)											
depositscaledXurban		−0.0043**										
		(−2.401)										
depositlengthXurban			−0.0026									
			(−1.646)									
depositmatchXurban				−0.0055***								
				(−3.357)								
depositcntXweightedpce					−0.0000**							
					(−2.265)							
depositscaledXweightedpce						−0.0000						
						(−1.257)						
depositlengthXweightedpce							−0.0000***					
							(−3.456)					
depositmatchXweightedpce								−0.0000**				
								(−2.030)				
depositcntXweightedinduscontrib									−0.0016			
									(−1.343)			
depositscaledXweightedinduscontrib										−0.0020		
										(−1.610)		
depositlengthXweightedinduscontrib											−0.0027	
											(−1.629)	
depositmatchXweightedinduscontrib												−0.0038***
												(−2.755)
depositcnt	0.0012				0.0051				−0.0012			
	(0.691)				(1.544)				(−1.618)			
depositscaled		0.0026*				0.0030				−0.0003		
		(1.740)				(0.830)				(−0.371)		
depositlength			−0.0003				0.0075**				−0.0016*	
			(−0.193)				(2.409)				(−1.875)	
depositmatch				0.0028*				0.0044				−0.0004
				(1.824)				(1.219)				(−0.487)
urban	−0.0020	−0.0025	−0.0018	−0.0028								
	(−0.260)	(−0.332)	(−0.241)	(−0.373)								
weightedpce					0.0000	−0.0000	0.0000	−0.0000				
					(0.083)	(−0.031)	(0.097)	(−0.007)				
weightedinduscontrib									0.0047**	0.0050***	0.0040**	0.0045**
									(2.550)	(2.651)	(2.108)	(2.376)
roa	−0.0212	−0.0215	−0.0226	−0.0199	−0.0357	−0.0340	−0.0323	−0.0320	−0.0264	−0.0250	−0.0284	−0.0250
	(−0.141)	(−0.143)	(−0.145)	(−0.128)	(−0.231)	(−0.218)	(−0.204)	(−0.200)	(−0.172)	(−0.162)	(−0.179)	(−0.157)
logassets	−0.0225***	−0.0227***	−0.0223***	−0.0229***	−0.0235***	−0.0236***	−0.0235***	−0.0239***	−0.0233***	−0.0236***	−0.0231***	−0.0238***
	(−6.105)	(−6.090)	(−6.134)	(−6.140)	(−6.496)	(−6.282)	(−6.569)	(−6.340)	(−6.155)	(−6.140)	(−6.172)	(−6.196)
nim	0.4684	0.4649	0.5078	0.4864	0.4725	0.4516	0.4978	0.4260	0.4422	0.4465	0.4635	0.4266
	(0.983)	(0.968)	(1.064)	(1.014)	(0.979)	(0.923)	(1.034)	(0.872)	(0.934)	(0.938)	(0.973)	(0.890)
loanlossprovisionratio	−0.5460**	−0.5529**	−0.5622**	−0.5667**	−0.5382**	−0.5485**	−0.5359**	−0.5523**	−0.5107**	−0.5155**	−0.5178**	−0.5129**
	(−2.166)	(−2.193)	(−2.235)	(−2.246)	(−2.132)	(−2.145)	(−2.134)	(−2.154)	(−1.999)	(−2.007)	(−2.034)	(−1.991)
ppop_loans	−0.0160***	−0.0157***	−0.0156***	−0.0154***	−0.0165***	−0.0162***	−0.0161***	−0.0159***	−0.0164***	−0.0162***	−0.0160***	−0.0162***
	(−3.622)	(−3.521)	(−3.563)	(−3.504)	(−3.725)	(−3.604)	(−3.632)	(−3.499)	(−3.577)	(−3.499)	(−3.508)	(−3.458)
debttoequity	−0.0004	−0.0004	−0.0003	−0.0004	−0.0004	−0.0004	−0.0004	−0.0004	−0.0003	−0.0003	−0.0003	−0.0003
	(−0.847)	(−0.888)	(−0.820)	(−0.884)	(−0.766)	(−0.813)	(−0.755)	(−0.897)	(−0.716)	(−0.718)	(−0.732)	(−0.730)
loantodepositratio	0.0234**	0.0235**	0.0235**	0.0237**	0.0239*	0.0251**	0.0231*	0.0253**	0.0249**	0.0248**	0.0250**	0.0252**
	(2.100)	(2.090)	(2.137)	(2.130)	(1.922)	(2.006)	(1.894)	(2.032)	(2.118)	(2.095)	(2.166)	(2.150)
coredeposits_assets	0.0100	0.0104	0.0097	0.0101	0.0091	0.0097	0.0088	0.0094	0.0095	0.0097	0.0092	0.0094
	(0.939)	(0.976)	(0.926)	(0.958)	(0.800)	(0.837)	(0.790)	(0.824)	(0.889)	(0.905)	(0.876)	(0.884)
dln_deposits_4q	0.0169	0.0174	0.0171	0.0176	0.0185	0.0192	0.0178	0.0191	0.0173	0.0176	0.0171	0.0171
	(0.819)	(0.837)	(0.833)	(0.848)	(0.888)	(0.913)	(0.864)	(0.910)	(0.832)	(0.842)	(0.830)	(0.825)
Constant	0.3604***	0.3653***	0.3577***	0.3673***	0.3732***	0.3777***	0.3747***	0.3818***	0.3689***	0.3748***	0.3662***	0.3780***
	(5.494)	(5.508)	(5.496)	(5.558)	(4.889)	(4.818)	(4.985)	(4.905)	(5.635)	(5.643)	(5.639)	(5.702)
Observations	8,276	8,276	8,276	8,276	7,737	7,737	7,737	7,737	8,220	8,220	8,220	8,220
Adjusted R-squared	0.117	0.117	0.118	0.118	0.119	0.119	0.120	0.119	0.118	0.118	0.119	0.119

Table 8: Interaction with banks' depositor characteristics. *Notes:* In Panel A, the variable of interest is the interaction between the deposit discussion metric and the bank's average urbanness. In Panel B, interest is on the interaction between deposit discussion and the bank's average personal consumption expenditure. In Panel C, the interest is on the interaction between deposit discussion and the contribution of the information sector to the local economy that the bank serves.

depositcntXrevmix	-0.0020^{***}			
	(-2.651)			
depositscaledXrevmix		-0.0040^{***}		
		(-2.997)		
depositlengthXrevmix			-0.0003	
			(-0.351)	
depositmatchXrevmix				-0.0006
				(-0.944)
depositcnt	-0.0010			
	(-1.417)			
depositscaled		0.0002		
		(0.275)		
depositlength			-0.0024^{***}	
			(-2.717)	
depositmatch				-0.0017^{**}
				(-2.031)
revmix	-0.0018	-0.0024^{**}	-0.0017	-0.0020
	(-1.283)	(-2.123)	(-1.190)	(-1.374)
roa	0.0671	0.1551	0.0376	0.0407
	(0.548)	(1.388)	(0.283)	(0.306)
logassets	-0.0226^{***}	-0.0231^{***}	-0.0223^{***}	-0.0230^{***}
	(-6.096)	(-6.135)	(-6.110)	(-6.125)
nim	0.3715	0.4136	0.3965	0.3818
	(0.759)	(0.831)	(0.816)	(0.787)
loanlossprovisionratio	-0.4787^{*}	-0.4219	-0.5177^{**}	-0.5193^{**}
	(-1.903)	(-1.628)	(-2.066)	(-2.060)
ppop_loans	-0.0158^{***}	-0.0158^{***}	-0.0151^{***}	-0.0153^{***}
	(-3.435)	(-3.446)	(-3.182)	(-3.248)
debttoequity	-0.0003	-0.0003	-0.0003	-0.0003
	(-0.843)	(-0.728)	(-0.836)	(-0.830)
loantodepositratio	0.0240^{**}	0.0241^{**}	0.0238^{**}	0.0241^{**}
	(2.136)	(2.098)	(2.162)	(2.166)
coredeposits_assets	0.0097	0.0100	0.0096	0.0100
	(0.911)	(0.945)	(0.916)	(0.946)
dln_deposits_4q	0.0169	0.0180	0.0170	0.0175
	(0.816)	(0.859)	(0.826)	(0.842)
Constant	0.3614^{***}	0.3687^{***}	0.3572^{***}	0.3679^{***}
	(5.594)	(5.639)	(5.585)	(5.638)
Observations	8,276	8,276	8,276	8,276
Adjusted R-squared	0.118	0.119	0.118	0.117

Table 9: Interaction with the extent of revenue mix. *Notes:* The deposit discussion metric is interacted with *revmix* defined as noninterest revenue divided by interest revenue.

	Panel A				Panel B			
deposityield	−0.0455	−0.0713	−0.1353	−0.1773	−0.1290	−0.1356	−0.1271	−0.1490
	(−0.069)	(−0.108)	(−0.206)	(−0.269)	(−0.191)	(−0.199)	(−0.194)	(−0.225)
depositcntXdeposityield					0.2701			
					(1.040)			
depositscaledXdeposityield						0.2310		
						(0.797)		
depositlengthXdeposityield							0.1948	
							(0.829)	
depositmatchXdeposityield								0.1501
								(0.555)
depositcnt	−0.0018***				−0.0025***			
	(−2.784)				(−2.711)			
depositscaled		−0.0010				−0.0016		
		(−1.498)				(−1.606)		
depositlength			−0.0026***				−0.0030***	
			(−3.187)				(−2.842)	
depositmatch				−0.0019**				−0.0022**
				(−2.350)				(−1.995)
roa	−0.0219	−0.0229	−0.0245	−0.0239	−0.0196	−0.0203	−0.0274	−0.0256
	(−0.144)	(−0.151)	(−0.157)	(−0.154)	(−0.129)	(−0.134)	(−0.176)	(−0.165)
logassets	−0.0224***	−0.0228***	−0.0223***	−0.0229***	−0.0225***	−0.0229***	−0.0222***	−0.0230***
	(−6.042)	(−6.035)	(−6.070)	(−6.084)	(−6.034)	(−6.066)	(−6.045)	(−6.076)
nim	0.4683	0.4777	0.5024	0.4836	0.4684	0.5093	0.5036	0.5134
	(0.990)	(1.004)	(1.060)	(1.016)	(0.990)	(1.079)	(1.063)	(1.080)
loanlossprovisionratio	−0.5458**	−0.5514**	−0.5509**	−0.5542**	−0.5451**	−0.5401**	−0.5537**	−0.5464**
	(−2.067)	(−2.082)	(−2.096)	(−2.097)	(−2.068)	(−2.025)	(−2.107)	(−2.062)
ppop_loans	−0.0160***	−0.0158***	−0.0155***	−0.0156***	−0.0161***	−0.0159***	−0.0155***	−0.0156***
	(−3.619)	(−3.558)	(−3.562)	(−3.538)	(−3.674)	(−3.618)	(−3.572)	(−3.554)
debttoequity	−0.0004	−0.0004	−0.0004	−0.0004	−0.0003	−0.0003	−0.0004	−0.0003
	(−0.869)	(−0.871)	(−0.854)	(−0.860)	(−0.826)	(−0.800)	(−0.856)	(−0.825)
loantodepositratio	0.0236**	0.0235**	0.0237**	0.0239**	0.0240**	0.0237**	0.0237**	0.0238**
	(2.107)	(2.093)	(2.141)	(2.141)	(2.127)	(2.106)	(2.112)	(2.129)
coredeposits_assets	0.0103	0.0107	0.0102	0.0106	0.0104	0.0106	0.0102	0.0105
	(0.966)	(0.994)	(0.960)	(0.994)	(0.975)	(0.990)	(0.963)	(0.989)
dln_deposits_4q	0.0172	0.0177	0.0175	0.0179	0.0171	0.0176	0.0173	0.0177
	(0.835)	(0.852)	(0.849)	(0.861)	(0.828)	(0.848)	(0.840)	(0.856)
Constant	0.3576***	0.3631***	0.3553***	0.3659***	0.3579***	0.3647***	0.3547***	0.3661***
	(5.553)	(5.566)	(5.559)	(5.611)	(5.539)	(5.592)	(5.535)	(5.607)
Observations	8, 276	8, 276	8, 276	8, 276	8, 276	8, 276	8, 276	8, 276
Adjusted R-squared	0.117	0.117	0.118	0.117	0.117	0.117	0.118	0.117

Table 10: Controlling for deposit rates. *Notes:* In Panel A, in the regression deposit rates are inserted as a control variable. In Panel B the interaction between deposit rates and our deposit discussion metrics is included.

	Panel A				Panel B				
	Dep.: insured time deposit rates next period				Dep.: insured time deposit flows next period				
depositYield_smalltime_p1					0.5770***	0.5950***	0.5822***	0.5803***	0.6369***
					(2.865)	(2.916)	(2.931)	(2.947)	(3.489)
depositcntXdepositYield_smalltime_p1						−0.0345			
						(−0.572)			
depositscaledXdepositYield_smalltime_p1							−0.0054		
							(−0.061)		
depositlengthXdepositYield_smalltime_p1								0.0165	
								(0.230)	
depositmatchXdepositYield_smalltime_p1									0.1830**
									(2.569)
depositcnt	0.0002***					−0.0002			
	(3.459)					(−0.482)			
depositscaled		0.0001**					−0.0002		
		(2.266)					(−0.345)		
depositlength			0.0000					−0.0004	
			(0.615)					(−0.876)	
depositmatch				−0.0001					−0.0008*
				(−1.352)					(−1.691)
roa	0.0061	0.0062	0.0062	0.0066	0.1033	0.1030	0.1032	0.1039	0.1038
	(0.392)	(0.396)	(0.399)	(0.425)	(0.939)	(0.935)	(0.929)	(0.942)	(0.946)
logassets	0.0002	0.0003	0.0003	0.0003	−0.0069***	−0.0068***	−0.0069***	−0.0069***	−0.0070***
	(1.149)	(1.325)	(1.274)	(1.298)	(−5.226)	(−5.244)	(−5.265)	(−5.230)	(−5.224)
nim	−0.0410	−0.0408	−0.0441	−0.0452	−0.6665**	−0.6680**	−0.6714**	−0.6693**	−0.6417**
	(−0.696)	(−0.691)	(−0.750)	(−0.771)	(−2.094)	(−2.106)	(−2.098)	(−2.102)	(−2.011)
loanlossprovisionratio	−0.0180	−0.0173	−0.0170	−0.0175	0.0925	0.0950	0.0927	0.0906	0.0936
	(−0.730)	(−0.704)	(−0.687)	(−0.713)	(0.407)	(0.418)	(0.406)	(0.396)	(0.413)
ppop_loans	−0.0004	−0.0004	−0.0005*	−0.0005*	−0.0028	−0.0028	−0.0028	−0.0027	−0.0029
	(−1.570)	(−1.621)	(−1.667)	(−1.654)	(−1.507)	(−1.534)	(−1.543)	(−1.503)	(−1.557)
debttoequity	0.0000	0.0000	0.0000	0.0000	−0.0001	−0.0001	−0.0001	−0.0001	−0.0001
	(0.174)	(0.178)	(0.183)	(0.201)	(−0.868)	(−0.860)	(−0.866)	(−0.850)	(−0.774)
loantodepositratio	0.0000	0.0000	0.0001	0.0001	0.0008	0.0008	0.0009	0.0010	0.0007
	(0.022)	(0.027)	(0.091)	(0.184)	(0.148)	(0.149)	(0.161)	(0.182)	(0.126)
coredeposits_assets	0.0001	0.0001	0.0001	0.0001	−0.0248***	−0.0249***	−0.0248***	−0.0248***	−0.0248***
	(0.215)	(0.174)	(0.179)	(0.187)	(−5.419)	(−5.435)	(−5.412)	(−5.433)	(−5.422)
dln_deposits_4q	−0.0008	−0.0009	−0.0009	−0.0009	0.0477***	0.0475***	0.0476***	0.0476***	0.0474***
	(−0.994)	(−1.060)	(−1.099)	(−1.122)	(5.663)	(5.635)	(5.638)	(5.636)	(5.647)
Constant	0.0005	−0.0000	0.0000	−0.0000	0.1339***	0.1325***	0.1337***	0.1326***	0.1346***
	(0.144)	(−0.010)	(0.013)	(−0.001)	(5.814)	(5.836)	(5.884)	(5.835)	(5.786)
Observations	8,177	8,177	8,177	8,177	8,177	8,177	8,177	8,177	8,177
Adjusted R-squared	0.765	0.764	0.764	0.764	0.152	0.152	0.152	0.152	0.152

Table 11: Insured time deposit rates. *Notes:* In Panel A, the dependent variable is the insured time deposit rates observed at $t+1$. In Panel B, the dependent variable is the change in insured time deposits at $t+1$. In columns 6 through 9, the interaction between deposit discussion and insured time deposit rates is introduced.

depositcntXi_daysuntilquarterend	-0.0030^{***}			
	(-3.076)			
depositscaledXi_daysuntilquarterend		-0.0028^{***}		
		(-2.715)		
depositlengthXi_daysuntilquarterend			-0.0024^{**}	
			(-2.379)	
depositmatchXi_daysuntilquarterend				-0.0030^{**}
				(-2.483)
depositcnt	-0.0002			
	(-0.203)			
depositscaled		0.0003		
		(0.355)		
depositlength			-0.0012	
			(-1.141)	
depositmatch				-0.0004
				(-0.430)
$\mathrm{i}_{d}aysuntilquarterend$	0.0026^{**}	0.0029^{**}	0.0026^{**}	0.0029^{**}
	(1.978)	(2.145)	(2.009)	(2.167)
roa	0.0420	0.0420	0.0359	0.0413
	(0.283)	(0.282)	(0.239)	(0.277)
logassets	-0.0192^{***}	-0.0196^{***}	-0.0191^{***}	-0.0198^{***}
	(-5.731)	(-5.730)	(-5.754)	(-5.823)
nim	0.4839	0.4885	0.5142	0.4997
	(0.996)	(1.005)	(1.057)	(1.028)
loanlossprovisionratio	-0.6475^{***}	-0.6579^{***}	-0.6633^{***}	-0.6619^{***}
	(-2.718)	(-2.737)	(-2.773)	(-2.754)
ppop_loans	-0.0145^{***}	-0.0142^{***}	-0.0142^{***}	-0.0142^{***}
	(-3.313)	(-3.248)	(-3.243)	(-3.233)
debttoequity	-0.0003	-0.0002	-0.0003	-0.0002
	(-0.615)	(-0.578)	(-0.648)	(-0.584)
loantodepositratio	0.0224^{**}	0.0225^{**}	0.0223^{**}	0.0225^{**}
	(2.536)	(2.539)	(2.574)	(2.591)
coredeposits_assets	0.0113	0.0113	0.0114	0.0110
	(1.086)	(1.097)	(1.100)	(1.072)
dln_deposits_4q	0.0049	0.0057	0.0054	0.0060
	(0.246)	(0.280)	(0.272)	(0.296)
Constant	0.3032^{***}	0.3079^{***}	0.3016^{***}	0.3128^{***}
	(5.108)	(5.132)	(5.103)	(5.215)
Observations	$8,359$	$8,359$	$8,359$	$8,359$
Adjusted R-squared	0.124	0.124	0.124	0.124

Table 12: Testing for heterogeneity depending on time between the date of earnings call and the end of the quarter. *Notes:* The deposit discussion metric is interacted with a dummy variable for whether or not the number of days between the date of the earnings call and the end of the quarter is above the median for the given quarter.

	Panel A				Panel B			
uninsuredfraction	−0.0885***	−0.0882***	−0.0870***	−0.0872***	−0.0914***	−0.0910***	−0.0885***	−0.0891***
	(−6.706)	(−6.714)	(−6.618)	(−6.648)	(−6.850)	(−6.810)	(−6.709)	(−6.720)
depositcntXuninsuredfraction					−0.0165***			
					(−3.407)			
depositscaledXuninsuredfraction						−0.0113**		
						(−2.041)		
depositlengthXuninsuredfraction							−0.0118**	
							(−2.165)	
depositmatchXuninsuredfraction								−0.0080
								(−1.640)
depositcnt	−0.0019***				0.0066***			
	(−2.772)				(2.651)			
depositscaled		−0.0010				0.0046*		
		(−1.309)				(1.750)		
depositlength			−0.0021**				0.0039	
			(−2.374)				(1.425)	
depositmatch				−0.0012				0.0027
				(−1.390)				(1.108)
roa	−0.0180	−0.0191	−0.0201	−0.0199	−0.0125	−0.0128	−0.0136	−0.0168
	(−0.096)	(−0.102)	(−0.105)	(−0.104)	(−0.067)	(−0.070)	(−0.072)	(−0.089)
logassets	−0.0191***	−0.0195***	−0.0192***	−0.0196***	−0.0197***	−0.0197***	−0.0193***	−0.0197***
	(−4.976)	(−4.978)	(−5.014)	(−4.982)	(−5.153)	(−5.039)	(−5.083)	(−4.992)
nim	0.4523	0.4661	0.4902	0.4801	0.3815	0.4459	0.4007	0.4393
	(0.951)	(0.974)	(1.030)	(1.004)	(0.811)	(0.938)	(0.824)	(0.902)
loanlossprovisionratio	−0.5809**	−0.5898**	−0.5944**	−0.5991**	−0.5597**	−0.5884**	−0.5639**	−0.5853**
	(−2.316)	(−2.344)	(−2.357)	(−2.369)	(−2.295)	(−2.365)	(−2.243)	(−2.304)
ppop_loans	−0.0019	−0.0017	−0.0017	−0.0017	0.0003	0.0002	−0.0019	−0.0017
	(−0.395)	(−0.356)	(−0.351)	(−0.352)	(0.065)	(0.051)	(−0.389)	(−0.348)
debttoequity	0.0000	−0.0000	0.0000	−0.0000	−0.0000	−0.0001	0.0000	−0.0000
	(0.007)	(−0.001)	(0.002)	(−0.002)	(−0.031)	(−0.159)	(0.039)	(−0.107)
loantodepositratio	0.0308***	0.0306***	0.0306***	0.0307***	0.0306***	0.0299***	0.0311***	0.0304***
	(3.470)	(3.442)	(3.466)	(3.446)	(3.373)	(3.254)	(3.646)	(3.418)
coredeposits_assets	0.0137	0.0140	0.0135	0.0138	0.0130	0.0129	0.0127	0.0127
	(1.263)	(1.286)	(1.251)	(1.274)	(1.211)	(1.197)	(1.193)	(1.183)
dln_deposits_4q	0.0177	0.0182	0.0180	0.0184	0.0174	0.0177	0.0176	0.0176
	(0.892)	(0.911)	(0.910)	(0.920)	(0.887)	(0.890)	(0.898)	(0.888)
Constant	0.3343***	0.3402***	0.3344***	0.3424***	0.3466***	0.3479***	0.3383***	0.3457***
	(4.968)	(4.982)	(4.987)	(4.989)	(5.219)	(5.121)	(5.090)	(5.054)
Observations	8,276	8,276	8,276	8,276	8,276	8,276	8,276	8,276
Adjusted R-squared	0.130	0.129	0.130	0.129	0.132	0.130	0.131	0.130

Table 13: Controlling for fraction uninsured. *Notes:* In Panel A, the uninsured fraction, measured as the uninsured deposit balance divided by the total deposit balance, is controlled for in the regression. In Panel B, the interaction between the deposit discussion metric and uninsured fraction is introduced.

	Panel A				Panel B			
uninsuredperperson	-0.0000^{**}	-0.0000^{***}	-0.0000^{***}	-0.0000^{***}	-0.0000^{*}	-0.0000	-0.0000^{**}	-0.0000^{**}
	(-2.588)	(-2.601)	(-2.665)	(-2.676)	(-1.908)	(-1.535)	(-2.561)	(-2.292)
depositcntXuninsuredperperson					-0.0000			
					(-1.306)			
depositscaledXuninsuredperperson						-0.0000		
						(-1.221)		
depositlengthXuninsuredperperson							-0.0000^{*}	
							(-1.806)	
depositmatchXuninsuredperperson								-0.0000^{*}
								(-1.773)
depositcnt	-0.0018^{***}				-0.0017^{**}			
	(-2.781)				(-2.593)			
depositscaled		-0.0010				-0.0009		
		(-1.484)				(-1.320)		
depositlength			-0.0026^{***}				-0.0025^{***}	
			(-3.173)				(-2.993)	
depositmatch				-0.0019^{**}				-0.0017^{**}
				(-2.317)				(-2.109)
roa	-0.0177	-0.0184	-0.0190	-0.0179	-0.0166	-0.0160	-0.0189	-0.0165
	(-0.116)	(-0.121)	(-0.121)	(-0.115)	(-0.109)	(-0.105)	(-0.121)	(-0.106)
logassets	-0.0223^{***}	-0.0226^{***}	-0.0222^{***}	-0.0228^{***}	-0.0223^{***}	-0.0225^{***}	-0.0221^{***}	-0.0225^{***}
	(-5.972)	(-5.965)	(-6.002)	(-6.010)	(-5.996)	(-5.901)	(-5.944)	(-5.732)
nim	0.5232	0.5320	0.5550	0.5351	0.5419	0.5716	0.5606	0.5667
	(1.113)	(1.125)	(1.181)	(1.134)	(1.164)	(1.231)	(1.200)	(1.222)
loanlossprovisionratio	-0.5475^{**}	-0.5556^{**}	-0.5609^{**}	-0.5681^{**}	-0.5488^{**}	-0.5613^{**}	-0.5616^{**}	-0.5752^{**}
	(-2.174)	(-2.200)	(-2.233)	(-2.252)	(-2.183)	(-2.246)	(-2.240)	(-2.311)
ppop_loans	-0.0157^{***}	-0.0155^{***}	-0.0152^{***}	-0.0153^{***}	-0.0157^{***}	-0.0154^{***}	-0.0151^{***}	-0.0150^{***}
	(-3.534)	(-3.470)	(-3.465)	(-3.437)	(-3.534)	(-3.448)	(-3.452)	(-3.369)
debttoequity	-0.0004	-0.0004	-0.0003	-0.0003	-0.0004	-0.0004	-0.0003	-0.0003
	(-0.846)	(-0.849)	(-0.834)	(-0.841)	(-0.868)	(-0.884)	(-0.834)	(-0.847)
loantodepositratio	0.0240^{**}	0.0240^{**}	0.0241^{**}	0.0244^{**}	0.0237^{**}	0.0238^{**}	0.0243^{**}	0.0250^{**}
	(2.158)	(2.143)	(2.194)	(2.192)	(2.128)	(2.126)	(2.229)	(2.289)
coredeposits_assets	0.0100	0.0103	0.0097	0.0101	0.0100	0.0104	0.0099	0.0106
	(0.943)	(0.969)	(0.926)	(0.955)	(0.941)	(0.974)	(0.938)	(0.990)
dln_deposits_4q	0.0163	0.0168	0.0165	0.0169	0.0161	0.0163	0.0165	0.0167
	(0.799)	(0.816)	(0.811)	(0.823)	(0.791)	(0.799)	(0.811)	(0.817)
Constant	0.3545^{***}	0.3601^{***}	0.3524^{***}	0.3630^{***}	0.3551^{***}	0.3583^{***}	0.3509^{***}	0.3563^{***}
	(5.459)	(5.474)	(5.467)	(5.514)	(5.486)	(5.409)	(5.412)	(5.229)
Observations	8,276	8,276	8,276	8,276	8,276	8,276	8,276	8,276
Adjusted R-squared	0.117	0.117	0.118	0.117	0.118	0.118	0.118	0.118

Table 14: Controlling for average uninsured account balanes. *Notes:* In Panel A, the average uninsured balance, measured as the size of uninsured deposits divided by the number of accounts that are uninsured, is controlled for in the regression. In Panel B, the interaction between the deposit discussion metric and the average uninsured balance is introduced.

	Panel A Presentation session only				Panel B Q&A session only			
depositcnt	−0.0018				−0.0021			
	(−0.800)				(−1.138)			
depositscaled		−0.0012				−0.0020		
		(−0.664)				(−1.102)		
depositlength			−0.0082***				−0.0023	
			(−2.808)				(−1.348)	
depositmatch				−0.0090***				−0.0019
				(−2.852)				(−1.063)
roa	−1.6350	−1.6324	−1.6176	−1.6527	−1.7340	−1.7346	−1.7388	−1.7373
	(−1.092)	(−1.090)	(−1.088)	(−1.109)	(−1.091)	(−1.091)	(−1.095)	(−1.095)
logassets	−0.0529***	−0.0531***	−0.0525***	−0.0542***	−0.0531***	−0.0534***	−0.0529***	−0.0534***
	(−4.229)	(−4.206)	(−4.171)	(−4.365)	(−4.206)	(−4.225)	(−4.175)	(−4.230)
nim	8.5609	8.5580	8.5539	8.5287	7.9137	7.9030	7.9543	7.9383
	(1.342)	(1.342)	(1.350)	(1.349)	(1.278)	(1.274)	(1.287)	(1.283)
loanlossprovisionratio	−15.9992	−16.0073	−16.0144	−16.0861	−16.1632	−16.1679	−16.1704	−16.1759
	(−1.398)	(−1.399)	(−1.404)	(−1.409)	(−1.397)	(−1.398)	(−1.399)	(−1.399)
ppop_loans	−0.0032	−0.0030	−0.0048	−0.0041	−0.0036	−0.0036	−0.0026	−0.0027
	(−0.057)	(−0.053)	(−0.085)	(−0.073)	(−0.064)	(−0.063)	(−0.046)	(−0.048)
debttoequity	0.0016	0.0016	0.0017	0.0017	0.0015	0.0015	0.0015	0.0015
	(0.477)	(0.473)	(0.509)	(0.503)	(0.451)	(0.454)	(0.447)	(0.448)
loantodepositratio	0.0377	0.0378	0.0376	0.0408	0.0404	0.0406	0.0405	0.0405
	(1.122)	(1.124)	(1.126)	(1.201)	(1.182)	(1.188)	(1.191)	(1.194)
coredeposits_assets	−0.0577	−0.0577	−0.0581	−0.0574	−0.0588	−0.0582	−0.0587	−0.0584
	(−1.385)	(−1.384)	(−1.407)	(−1.377)	(−1.423)	(−1.411)	(−1.420)	(−1.417)
dln_deposits_4q	0.1229**	0.1232**	0.1209**	0.1214**	0.1253**	0.1258**	0.1261**	0.1267**
	(2.265)	(2.272)	(2.265)	(2.265)	(2.366)	(2.381)	(2.391)	(2.412)
Constant	0.8472***	0.8508***	0.8391***	0.8655***	0.8541***	0.8583***	0.8517***	0.8597***
	(3.321)	(3.308)	(3.269)	(3.412)	(3.333)	(3.349)	(3.308)	(3.353)
Observations	8,276	8,276	8,276	8,276	8,252	8,252	8,252	8,252
Adjusted R-squared	0.131	0.131	0.132	0.133	0.132	0.132	0.132	0.132

Table 15: Presentations vs Q&A. *Notes:* In Panel A, the deposit discussion metric is constructed on only based on the text from the presentation session of the earnings call. In Panel B, the deposit discussion metric is constructed only based on the text from the Q&A session of the earnings call.

Appendix

	Dep: insured deposit change			
depositcntXiave_negsentiment	−0.0001			
	(−0.210)			
depositscaledXiave_negsentiment		−0.0005		
		(−0.581)		
depositlengthXiave_negsentiment			−0.0005	
			(−0.587)	
depositmatchXiave_negsentiment				−0.0011
				(−1.115)
depositcnt	0.0011			
	(1.335)			
depositscaled		0.0018**		
		(1.977)		
depositlength			0.0003	
			(0.362)	
depositmatch				0.0012
				(1.165)
iave_negsentiment	−0.0010	−0.0010	−0.0010	−0.0011
	(−0.834)	(−0.828)	(−0.842)	(−0.882)
roa	−0.2047	−0.2069	−0.2022	−0.2005
	(−0.575)	(−0.583)	(−0.568)	(−0.565)
logassets	−0.0236***	−0.0235***	−0.0235***	−0.0234***
	(−6.757)	(−6.788)	(−6.766)	(−6.759)
nim	1.0863	1.1159	1.0580	1.0692
	(1.606)	(1.648)	(1.559)	(1.576)
loanlossprovisionratio	−2.1614	−2.1622	−2.1558	−2.1555
	(−1.269)	(−1.271)	(−1.266)	(−1.264)
ppop_loans	−0.0168***	−0.0168***	−0.0170***	−0.0171***
	(−3.346)	(−3.358)	(−3.389)	(−3.411)
debttoequity	−0.0003	−0.0003	−0.0003	−0.0003
	(−0.906)	(−0.913)	(−0.896)	(−0.900)
loantodepositratio	0.0360***	0.0357***	0.0363***	0.0361***
	(5.085)	(5.081)	(5.059)	(5.028)
coredeposits_assets	0.0095	0.0093	0.0094	0.0094
	(1.056)	(1.039)	(1.035)	(1.042)
dln_deposits_4q	0.0686**	0.0686**	0.0683**	0.0683**
	(2.268)	(2.267)	(2.250)	(2.253)
Constant	0.3619***	0.3592***	0.3589***	0.3581***
	(6.162)	(6.175)	(6.165)	(6.150)
Observations	8,339	8,339	8,339	8,339
Adjusted R-squared	0.095	0.095	0.095	0.095

Table A1: Interaction with sentiment for predicting insured deposit flows. *Notes:* This table repeats the analysis shown in Table 5. The dependent variable is the change in the insured deposits measured from t to t+1.

	Panel A				Panel B				Panel C			
depositcntXurban	0.0028											
	(1.134)											
depositscaledXurban		0.0015										
		(0.733)										
depositlengthXurban			0.0008									
			(0.427)									
depositmatchXurban				−0.0006								
				(−0.250)								
depositcntXweightedpce					0.0000*							
					(1.723)							
depositscaledXweightedpce						0.0000						
						(0.680)						
depositlengthXweightedpce							0.0000**					
							(2.333)					
depositmatchXweightedpce								0.0000				
								(1.229)				
depositcntXweightedinduscontrib									−0.0005			
									(−0.385)			
depositscaledXweightedinduscontrib										−0.0025**		
										(−2.438)		
depositlengthXweightedinduscontrib											0.0017	
											(1.132)	
depositmatchXweightedinduscontrib												−0.0008
												(−0.641)
depositcnt	−0.0014				−0.0041				0.0013			
	(−0.665)				(−1.330)				(1.456)			
depositscaled		0.0002				−0.0004				0.0024***		
		(0.119)				(−0.127)				(2.970)		
depositlength			−0.0006				−0.0076**				−0.0005	
			(−0.376)				(−2.320)				(−0.523)	
depositmatch				0.0010				−0.0033				0.0009
				(0.490)				(−0.972)				(0.967)
urban	−0.0156*	−0.0152	−0.0153*	−0.0152*								
	(−1.681)	(−1.650)	(−1.656)	(−1.661)								
weightedpce					0.0000	0.0000	0.0000	0.0000				
					(0.969)	(1.053)	(0.992)	(1.055)				
weightedinduscontrib									0.0013	0.0012	0.0020	0.0012
									(0.473)	(0.435)	(0.745)	(0.463)
roa	−0.2020	−0.2049	−0.1999	−0.2010	−0.1973	−0.2023	−0.1979	−0.1993	−0.2045	−0.2051	−0.2022	−0.2033
	(−0.567)	(−0.577)	(−0.561)	(−0.564)	(−0.540)	(−0.555)	(−0.542)	(−0.545)	(−0.572)	(−0.576)	(−0.565)	(−0.568)
logassets	−0.0238***	−0.0236***	−0.0236***	−0.0235***	−0.0254***	−0.0254***	−0.0250***	−0.0253***	−0.0246***	−0.0244***	−0.0244***	−0.0243***
	(−6.890)	(−6.933)	(−6.906)	(−6.910)	(−6.669)	(−6.762)	(−6.630)	(−6.714)	(−6.937)	(−6.949)	(−6.925)	(−6.939)
nim	1.0699	1.1061	1.0426	1.0585	1.2287*	1.2694*	1.2097*	1.2484*	1.1300*	1.1431*	1.1175	1.1055
	(1.572)	(1.626)	(1.527)	(1.552)	(1.703)	(1.771)	(1.673)	(1.733)	(1.657)	(1.675)	(1.630)	(1.610)
loanlossprovisionratio	−2.1714	−2.1697	−2.1607	−2.1577	−2.2059	−2.2038	−2.2099	−2.1985	−2.1602	−2.1574	−2.1554	−2.1456
	(−1.285)	(−1.284)	(−1.279)	(−1.275)	(−1.269)	(−1.269)	(−1.273)	(−1.265)	(−1.265)	(−1.266)	(−1.263)	(−1.254)
ppop_loans	−0.0165***	−0.0167***	−0.0168***	−0.0168***	−0.0171***	−0.0171***	−0.0171***	−0.0174***	−0.0172***	−0.0171***	−0.0174***	−0.0174***
	(−3.313)	(−3.343)	(−3.365)	(−3.383)	(−3.281)	(−3.312)	(−3.317)	(−3.390)	(−3.446)	(−3.391)	(−3.504)	(−3.486)
debttoequity	−0.0003	−0.0003	−0.0003	−0.0003	−0.0003	−0.0003	−0.0004	−0.0003	−0.0003	−0.0003	−0.0003	−0.0003
	(−0.894)	(−0.879)	(−0.880)	(−0.883)	(−0.854)	(−0.844)	(−0.904)	(−0.793)	(−0.831)	(−0.825)	(−0.868)	(−0.842)
loantodepositratio	0.0368***	0.0364***	0.0369***	0.0366***	0.0377***	0.0364***	0.0387***	0.0369***	0.0371***	0.0367***	0.0374***	0.0371***
	(5.187)	(5.178)	(5.139)	(5.101)	(4.878)	(4.889)	(4.846)	(4.871)	(5.175)	(5.168)	(5.157)	(5.123)
coredeposits_assets	0.0093	0.0089	0.0091	0.0090	0.0093	0.0087	0.0092	0.0089	0.0091	0.0084	0.0092	0.0090
	(1.032)	(0.995)	(0.995)	(0.989)	(0.957)	(0.891)	(0.953)	(0.915)	(1.002)	(0.938)	(1.016)	(0.990)
dln_deposits_4q	0.0687**	0.0687**	0.0684**	0.0684**	0.0683**	0.0683**	0.0684**	0.0680**	0.0691**	0.0686**	0.0692**	0.0687**
	(2.287)	(2.283)	(2.267)	(2.265)	(2.366)	(2.355)	(2.370)	(2.348)	(2.319)	(2.296)	(2.321)	(2.287)
Constant	0.3762***	0.3738***	0.3734***	0.3725***	0.3654***	0.3640***	0.3579***	0.3617***	0.3753***	0.3724***	0.3716***	0.3710***
	(6.401)	(6.414)	(6.413)	(6.393)	(5.449)	(5.500)	(5.387)	(5.470)	(6.286)	(6.292)	(6.271)	(6.273)
Observations	8,342	8,342	8,342	8,342	7,800	7,800	7,800	7,800	8,286	8,286	8,286	8,286
Adjusted R-squared	0.096	0.096	0.095	0.096	0.094	0.094	0.094	0.094	0.096	0.096	0.096	0.096

Table A2: Interaction with banks' depositor characteristics for predicting insured deposit flows. *Notes:* This table repeats the analysis shown in Table 8. The dependent variable is the change in the insured deposits measured from t to t+1.