# New Approaches in Financial Economics: An Exposition using Financial Markets and Banking Regulation

by

Katherine Wood

(Under the Direction of Annette Poulsen)

Abstract

This dissertation comprises an introductory chapter as well as two empirical studies on new approaches to questions in financial economics. In the first chapter, I examine the impact of two bank regulation changes on liquidity creation. Using amendments in 2005 to the FDIC Improvement Act (FDICIA) and the Community Reinvestment Act (CRA), I document that treating each regulatory change as a separate event leads to confounding results. Failure to disentangle the two changes leads to an overstatement of the increase in liquidity creation driven by regulatory changes. I find no evidence that the regulatory change leads to outcomes contradictory to the purposes of the CRA. The second chapter examines how uncertainty affects the rationality of the market's earnings expectations. The results reveal cross-sectional and time-series variations in uncertainty are associated with opposite patterns in the return predictability of analysts' conditional biases. An information choice model explains the contrasting findings and helps to reconcile the existing theories in the literature. The model suggests that for highly uncertain firms, investors find it too costly to acquire new information. However, when aggregate uncertainty increases temporarily, the benefit of acquiring new information outweighs the cost of acquiring such information.

Index words: FDICIA, CRA, Liquidity Creation, Regulatory Compliance, Rational Inattention, Volatility, Analyst Earnings Forecasts, Machine Learning

# New Approaches in Financial Economics: An Exposition using Financial Markets and Banking Regulation

by

Katherine Wood

BS, Kansas State University, 2014

MSc, University of Manchester, England, 2018

A Dissertation Submitted to the Graduate Faculty of the

University of Georgia in Partial Fulfillment of the Requirements for the Degree.

Doctor of Philosophy

Athens, Georgia

2023

# New Approaches in Financial Economics: An Exposition using Financial Markets and Banking Regulation

by

Katherine Wood

Major Professor: Annette Poulsen

Committee: Pekka Honkanen
Zhongjin Lu
Siddharth Vij

Electronic Version Approved:

Ron Walcott
Dean of the Graduate School
The University of Georgia
May 2023

# Dedication

To my family for their patience, understanding, and support throughout this entire process. Without you, none of this would have been possible. To Dr. Ansley Chua for encouraging me to pursue my degree and keeping me motivated along the way. Finally, to Forest and Wesley who keep me sane and make excellent supervisors.

# Acknowledgments

I would like to express my gratitude to my dissertation committee members—Annette Poulsen, Pekka Honkanen, Zhongjin Lu, and Siddharth Vij. Without their help, along with support from the rest of the Finance Department, my dissertation would not have been possible. I would also like to thank my fellow Ph.D. students at the University of Georgia for their friendship and encouragement during my time here.

# Contents

# List of Figures

# List of Tables

# Chapter 1

# Introduction

Financial economics encompasses a vast array of topics. Under the *Journal of Economic Literature* (JEL) classification system, financial economics consists of the following areas of study: financial crises, general financial markets, financial institutions and services, corporate finance and governance, behavioral finance, and household finance. On SSRN alone, there are currently 219,357 papers in the Financial Economics category, making it the fourth largest area of research in Social Sciences on SSRN.[1] From early papers such as Modigliani and Miller, 1958, researchers in financial economics have sought to explore and understand the mechanics that make up the financial market. As time has passed, researchers continue to revisit key questions in finance. This can be driven by new ideas, new data, advances in technology, etc. These allow researchers to take new approaches to address and investigate existing findings and theories. In this chapter, I introduce the two essays that comprise my dissertation and provide a brief foundation on their place in financial economics.

[1]This count was obtained on March 20, 2023.

## 1.1 Banking Regulations

Financial institutions play a critical role in the financial markets by providing liquidity and transforming risks (Bhattacharya & Thakor, 1993). Thus maintaining bank stability is a key concern for regulators. As an example, following the vast number of bank failures in the Great Depression, the Federal Deposit Insurance Company was created as part of The Banking Act of 1933 which established a system for deposit insurance.[2] Nobel prize winners Douglas Diamond and Philip Dybvig show in Diamond and Dybvig, 1983 how deposit insurance is critical in helping to prevent bank runs.

Banking regulations are not only enacted after crises but often are put into place to promote better stability and transparency. However, compliance with banking regulations can be costly for banks (Cyree, 2016; Dahl & Franke, 2017). Compliance costs can divert bank resources away from core functions such as lending. Often at the prompting of financial institutions, bank regulators sometimes identify situations where banks are unduly burdened by the cost of compliance with regulatory requirements. Often this excess regulatory burden most strongly impacts smaller banks. Accordingly, banking regulations are sometimes amended to help relieve some of the regulatory burden faced by banks. Researchers are able to utilize econometric tools to evaluate the impact and consequences of such regulatory changes.

Chapter two of my dissertation investigates two regulatory changes in 2005 aimed at reducing the regulatory burden for community banks: the Community Reinvestment Act

[2] See https://www.fdic.gov/bank/historical/firstfifty/ for a detailed history of the FDIC.

(CRA) and FDIC Improvement Act (FDICIA). The goal of the CRA is to encourage banks to meet the credit needs of the community, especially in low-income areas. As part of this, banks must submit detailed loan-level reports to evidence compliance. The FDICIA aims to ensure banks are adequately capitalized. The FDICIA requires a financial audit and assurance services for internal controls over financial reporting. I evaluate whether the amendment to these regulations lead to changes in bank behavior, specifically with respect to liquidity creation. If liquidity creation increases following a removal of regulatory requirements imposed on banks, it would suggest the regulatory change lead to a reduction in regulatory burden.

Prior literature has evaluated the impact of the FDICIA areas such as risk (Jin, Kanagaretnam, & Lobo, 2013; Jin, Kanagaretnam, Lobo, & Mathieu, 2013), CAMELS ratings (Gopalan et al., 2020), and allowances for loan and lease losses (Altamuro & Beatty, 2010). However, to my knowledge, research is largely silent on the CRA from a banking perspective, especially as it relates to the 2005 regulatory change.[3] I add to the literature by addressing both the CRA regulatory change as well as addressing an important methodological consideration that has not been addressed.

Before 2005, banks with total assets below $250 million dollars were exempt from generating the detailed CRA report. However, in 2005, the size-based threshold that exempted banks from submitting the loan level report was raised to $1 billion. Similarly, prior to 2005, banks with assets above $500 million were required to comply with the audit and assurance requirements of the FDICIA. However, following the amendment in 2005, banks with assets above $500 million but below $1 billion only had to obtain a financial audit and were exempt

[3]Much of the CRA literature evaluates the social impacts of the regulation.

from the assurance services requirement. I evaluate the impact of both regulatory changes by investigating whether changes occurred to liquidity creation for affected banks.

As previously discussed, liquidity creation is a key function of a bank. At a high level, this is the process by which banks take in liquid liabilities (i.e. deposits) and use them to finance illiquid assets (i.e. loans). The bank provides liquidity to credit-constrained individuals through lending while maintaining liquidity for depositors. To conduct this analysis, I utilize the liquidity creation measurement developed by Berger and Bouwman, 2009. This measure captures the liquidity creation process on both the asset and liability sides of the balance sheet from on- and off-balance sheet activity. As liquidity creation represents a key service provided by banks, high regulatory burdens may cause banks to reduce their liquidity creation due to costs of compliance. If this is the case, and if the regulatory burden of banks decreases following the regulatory amendments, liquidity creation should increase for affected banks.

I investigate the reduction in regulatory requirements of the CRA and FDICIA following the 2005 amendments. I provide novel evidence showing that treating each regulatory change that occurred as an independent event can lead to inaccurate inferences. This is because an overlapping set of banks were affected by the CRA and FDICIA regulation changes. My results show that failure to account for this overlap would cause researchers to conclude that both the CRA and the FDICIA regulatory changes lead to an increase in liquidity creation. However, after correctly identifying the CRA-specific and FDICIA-specific effects, I show that only the CRA regulatory change leads to an increase in liquidity creation.

As the goal of the CRA is to ensure the credit needs of the community in which the bank operates are met, especially in low-income areas, I further investigate the liquidity creation increase surrounding the CRA regulatory change. As small businesses are a key component of the community, I evaluate whether affected banks' lending decisions for small business lending change following the CRA amendment. It could be the case that even though overall bank lending increased, due to the removal of the regulatory oversight through the CRA report, banks may decide to lower their lending to small businesses. If banks view small business lending as inherently riskier than other types of lending, they may change their lending decisions following the removal of the CRA reporting requirement. I show that the overall dollar volume of small business lending increases following the CRA amendment. This suggests that the CRA regulatory change did not lead to actions by banks that were contrary to the goals of the CRA.

## 1.2 Market Rationality

One of the first topics in my introductory finance course has to do with the concept of market efficiency. Eugene Fama, in his well-known 1970 paper (Fama, 1970), provides an excellent summary of what it means for a market to be efficient:

> *In general terms, the ideal is a market in which prices provide accurate signals for resource allocation: that is, a market in which firms can make production-investment decisions, and investors can choose among the securities that represent ownership of firms' activities under the assumption that security prices at any*

> *time "fully reflect" all available in-formation. A market in which prices always "fully reflect" available information is called "efficient."*

At the heart of the efficient market hypothesis is the idea of rationality. For markets to be efficient, the market as a whole must be rational in its expectations. If market rationality holds, in situations where pricing is incorrect, investors should identify and take advantage of any arbitrage opportunities and thus restore prices to their correct levels. This would further suggest that in a rational market, there should be no evidence of predictability of future returns. However, the theory of market efficiency has been questioned due to numerous papers finding return predictability anomalies. The anomalies, or factors, identified are used to build models that explain documented return predictability. As an example, Feng et al., 2020 identifies and investigates 150 previously identified factors, which leads to the phrase the factor zoo.

Chapter 3 of my dissertation analyzes the degree to which uncertainty affects market rationality. The existing informational choice literature suggests that high uncertainty should lead to more rational market expectations due to an increase in the value of information associated with increased uncertainty (Baley & Veldkamp, 2021; Mackowiak et al., 2021). However, behavioral models suggest that when uncertainty is high, markets become less rational due to investors' behavioral biases (Hirshleifer, 2001).

We document that variations in time-series and cross-sectional uncertainty have opposite patterns in return predictability. We find that when periods of high time-series uncertainty, the return predictability of analysts' conditional biases is weaker. However, for firms that are more uncertain relative to the cross-sectional distribution of firms, the return predictability of analysts' conditional biases is stronger.

Using an information choice model that includes the information cost channel, we explain these opposite relationships. Investors must acquire information to unravel the conditional biases in analysts' forecasts. However, information acquisition is costly for the investor. When uncertainty is higher, the value of acquiring additional information is more beneficial. However, higher uncertainty means that less information is available, leading to higher information costs. Based on the model, firms with greater uncertainty relative to other firms in the cross-section will be more costly for investors to unravel the conditional biases in analysts' forecasts. This will lead to the documented positive relationship between cross-sectional uncertainty and the return predictability of analysts' conditional biases being stronger. However, when uncertainty is high in the time-series, for a given firm, the cost to acquire information about a firm is unlikely to spike due to a temporary increase in uncertainty. Thus, the model predicts a negative relationship between time-series uncertainty and the return predictability of analysts' conditional biases.

# Chapter 2

# A Tale of Two Banking Regulations

I examine how changes to a bank's regulatory requirements affect liquidity creation. Using amendments in 2005 to the FDIC Improvement Act (FDICIA) and the Community Reinvestment Act (CRA), I document that treating each regulatory change as a separate event leads to confounding results. Ignoring the overlap of regulations, the results suggest that both regulations lead to increases in liquidity creation. However, after disentangling the effects of the two regulatory changes, I show that only the amendment to the CRA drives the results. Failure to disentangle the two changes leads to an overstatement of the increase in liquidity creation. Furthermore, I find no evidence that the regulatory change leads to outcomes contradictory to the purposes of the CRA.

## 2.1 Introduction

Regulatory requirements based on size thresholds exist in all areas of banking. Size-based thresholds are situations where banks are required to comply with part or all of a regulation based on their total assets. [1] Changes in regulatory requirements across size-based thresholds have been shown to impact banks in areas such as M&A activity (Bindal et al., 2020), financial reporting (Altamuro & Beatty, 2010), examination ratings (Gopalan et al., 2020), etc. In this paper, I analyze the effects of regulatory changes with respect to liquidity creation, an essential function of a bank. Specifically, I examine community banks using 2005 amendments to the FDIC Improvement Act (FDICIA) and the Community Reinvestment Act (CRA). Both amendments reduced the regulatory requirements of affected banks. Affected banks were determined using size-based thresholds.[2] Changes to the FDICIA removed requirements related to the evaluation of internal controls over financial reporting (ICFR) while the CRA regulatory change removed certain loan reporting requirements. Specifically, I address the consequences caused by a failure to appropriately address the overlap in the banks affected by the CRA and FDICIA regulations. I show that failure to separate the CRA-specific and FDICIA-specific effects leads to erroneous conclusions about the changes to the two regulations using liquidity creation as a vehicle for the analysis.

[1]Some of the regulations with size-based thresholds are Dodd–Frank Wall Street Reform and Consumer Protection Act (Dodd–Frank), the FDIC Improvement Act (FDICIA), the Community Reinvestment Act (CRA), the Home Mortgage Disclosure Act (HMDA), etc.

[2]In 2020, the FDIC defined community banks as those with assets between \$500 million and \$1 billion. Smaller community banks are those with assets between \$100 million and \$500 million, and the smallest community banks are those with assets below \$100 million (Breitenstein et al., 2020).

Liquidity creation is an essential function of banks (Bryant, 1980; Diamond & Dybvig, 1983). Liquidity creation theory states that liquidity creation is the process by which banks transform illiquid assets (e.g. loans) into liquid liabilities (e.g. deposits). The bank then holds illiquid assets to provide the public with liquid assets. From the bank's perspective, the liquidity creation cycle begins with the bank financing illiquid assets with liquid liabilities. These illiquid assets are then transformed into further liquid liabilities. This process is enabled by the fact that banks are required to set aside only a fraction of their loan portfolio as a capital buffer. The remaining balance of the loan portfolio could, in theory, be completely converted into liquid liabilities, to be available to generate further loans.

Often liquidity creation is thought of from the lending side as previously discussed. However, Kashyap et al., 2002 show that banks provide liquidity through both their lending and deposit functions. Lending allows banks to provide borrowers with funds to make investments. At the same time, banks provide a safe location where depositors can keep and readily withdraw their funds. Banks thus generate liquidity for borrowers while maintaining liquidity for depositors. In addition, off-balance sheet liquidity creation occurs through the use of commitments (lines of credit) (Holmström & Tirole, 1998; Kashyap et al., 2002).

Liquidity creation has been shown to have a strong positive relationship with real economic output (Berger & Sedunov, 2017). In addition, Berger and Sedunov, 2017 show that liquidity creation is a better measure of bank output than loans or total assets. Measures of output such as loans or total assets do not consider the actions of the entire bank, whereas liquidity creation captures changes in assets and liabilities from on- and off-balance sheet activity. Much of the previous literature has focused on links between bank capital and liquidity creation (e.g.

Berger and Bouwman, 2009; Casu et al., 2019; Fungáčová et al., 2017). However, I examine liquidity creation from a different angle and provide empirical evidence of banks' reactions to changes in regulations surrounding size-based thresholds. The change in the regulations I consider, the CRA and FDICIA, have no direct impact on bank capital. In a similar paper, Nguyen et al., 2020 considers the impact of mandated stress tests for banks above a certain asset size. Stress testing, however, is inherently linked to bank capital. My analysis thus provides a new dynamic by considering the consequences of regulatory changes not directly linked to bank capital.

I investigate the effects of bank regulatory changes on liquidity creation. To do so, I use regulatory changes in the FDICIA and CRA that affect banks based on size thresholds. The FDICIA imposes greater regulatory requirements for banks above $500 million in assets. Above this level, prior to 2005, banks were required to conduct an external financial audit and provide an assessment by management and an independent accountant on the effectiveness of internal controls over financial reporting (ICFR). In 2005, banks below $1 billion but above $500 million in assets were exempted from needing to provide the assessments of ICFR.[3] Before 2005, the CRA required banks above $250 million to submit detailed, loan-level reports to evidence compliance in meeting the credit needs of the community, including low-income individuals. The regulatory change in 2005 increased the threshold of the reporting requirement to $1 billion, thus exempting banks with assets between $250 million and $1 billion from generating the reports.

[3]The FDICIA-compliant banks are still required to obtain a financial audit.

I have two main focuses in this paper. First, I show the consequences of analyzing the CRA and FDICIA regulatory changes independently. Banks with assets between $250 million and $500 million in assets are only affected by changes to the CRA regulation. However, banks between $500 million and $1 billion in assets are affected by both regulatory changes. This leads to an overlap in banks affected by the regulatory changes that is problematic for analysis. For example, if I analyze the FDICIA using banks above $500 million in assets as the treated sample and banks below this as the control sample, only part of the control group is affected by the CRA change (those with assets above $250 million). This leads to confounding results due to the overlap of banks affected by the two regulations. As liquidity creation is a main function of a bank, I use this as a laboratory in which to address this analysis.

Utilizing a differences-in-differences (DiD) framework, I first evaluate each regulatory change independently (i.e. I do not address the overlap of affected banks). In my analyses, banks that experience a change in their regulatory requirements from the FDICIA or CRA are designated as the treated group. These are banks with assets above $500 million and $250 million for the FDICIA and CRA, respectively. Banks below the given thresholds are designated as the control sample for each analysis. When I treat each regulation separately, I find evidence that suggests that banks respond to both the change in the FDICIA and CRA by increasing liquidity creation by 6.86% and 5.78%, respectively. I then compare these results to an analysis that correctly separates the CRA-specific and FDICIA-specific effects.

To do this, I set up two sub-samples. To analyze the effect of the CRA on liquidity creation, I exclude banks that are required to be compliant with the FDICIA, i.e. banks with assets above \$500 million. This means that my treated sample only consists of banks affected by the change in the CRA regulation. This allows me to cleanly identify the CRA-specific effect and provides a clean analysis of the changes to liquidity creation driven by the changes to the CRA. The construction of the sub-sample for the FDICIA is less straightforward as I need the variation in regulation between the control and treatment group to only come from the FDICIA. All banks affected by the change to the FDICIA are also affected by the change to the CRA. However, only part of the sample of banks unaffected by the FDICIA are affected by the CRA. To address this, I limit banks in the control sample to only include those affected by the CRA (i.e. banks above \$250 million in assets). By only using CRA-treated banks in the control sample for the FDICIA, I remove any CRA-specific effects. This allows me to evaluate the effect of each regulation without contamination from the other regulation. Following this update to my analysis, I find that only the CRA regulatory change leads to an increase in liquidity creation. My findings suggest that failure to address this overlap in the FDICIA and CRA regulations leads to confounding results that overstate liquidity creation.

Following the finding that only the CRA regulatory change leads to an increase in liquidity creation of 5.15%, I next investigate the effects of the CRA regulatory change in greater detail. The purpose of the CRA is to encourage banks to lend to the communities in which they operate, especially to low-income individuals. Even though I find that liquidity creation, and thus lending, increases for treated banks as a whole, it could be the case that the decrease in regulatory oversight leads to banks conducting less lending to small businesses, especially those in low-income areas, if banks consider these to be riskier loan types. If so, this would suggest that the regulatory change had outcomes contradictory to the goals of the CRA. I investigate changes to small business lending as well as loan-level analysis of SBA lending to provide some analysis to address this concern.

I develop three hypotheses addressing the impact of changes to bank regulatory requirements related to size thresholds for community banks.

> **Hypothesis 1:** Banks react to decreases in regulatory requirements by increasing their liquidity creation.
>
> **Hypothesis 1a:** Changes to liquidity creation are driven by the change to the CRA regulation.
>
> **Hypothesis 1b:** Changes to liquidity creation are driven by the change to the FDICIA regulation.

This is similar to the analysis in Berger et al., 2016. They find that regulatory intervention negatively impacts the activities of a bank. Instead of individual bank interventions, which can be thought of as increased regulatory oversight of an individual bank, I consider overall regulatory oversight changes. Regulations such as the FDICIA or CRA require affected banks to devote resources to maintain compliance. The resources cannot be used to elicit deposits or make loans, thus possibly reducing the liquidity creation of the bank. Similarly, a reduction in a bank's regulatory requirements can allow the reallocation of previously constrained resources. To test this hypothesis, I utilize the regulatory changes in 2005 to the FDICIA and CRA. When treated as independent events, thus ignoring the overlap in affected banks, the decrease in regulatory requirements from both regulations appears to lead to increases in liquidity creation.

Analyzing each regulation change independently, without further considerations, could lead to confounding results. Both the FDICIA and CRA regulatory changes occurred in the second half of 2005 and there is an overlap of the banks affected by both regulations. The FDICIA regulatory change affected banks with assets between \$500 million and \$1 billion, and the CRA change affected banks with assets between \$250 million and \$1 billion. Thus, there is possible contamination in the sample due to an overlap in banks based on the size thresholds of the two regulatory changes. To address this, I construct two sub-samples that allow me to cleanly identify the CRA-specific and/or FDICIA-specific results. To evaluate the impact of the CRA change, I exclude banks that are required to be compliant with the FDICIA, so banks above \$500 million in assets. This means that both my treated and control groups are not affected by changes to the FDICIA. For the FDICIA, all treated banks are also affected by the CRA change. However, if I used all banks below \$500 million in assets as the control sample, some banks are affected by the CRA change and some are not. As such, I limit the control sample to exclude banks not affected by the CRA change, removing any CRA-specific effects. Figure 2.1 provides a visual comparison of the analysis if each regulation was treated independently and my corrected samples to address the overlap.

Using the full sample, the results suggest that banks affected by the CRA and FDICIA regulatory changes increase their liquidity creation. However, when I use my preferred sub-samples to address concerns about the overlap between the regulations, I find that the results related to the changes in the FDICIA become insignificant, but the results continue to hold for the CRA regulation change. This provides evidence of the necessity to evaluate the effect of the CRA and FDICIA regulatory changes using uncontaminated/non-overlapping samples.

**Hypothesis 2:** The CRA regulatory change does not negatively impact a bank's actions to meet the credit needs of the community.

Previous results show that treated banks increase their liquidity creation following the 2005 CRA regulatory change. Further, I find that the increase in liquidity creation is driven by the standard liquidity creation process–an increase in both liquid liabilities and illiquid assets. This suggests that banks increase overall lending following the regulatory changes. However, the concern might be that because affected banks are no longer required to submit CRA loan reports, banks may decrease lending to the community, especially in low-income areas, while increasing lending in wealthier areas. If banks consider small businesses and low-income areas to be riskier, they may shift their resources to less risky loans following the removal of the loan-level reports. If this occurs, it suggests a negative outcome to the change in the regulation as it would be contrary to the goals of the CRA. Ideally, I would directly test if affected banks change their lending to their local community, especially in low-income neighborhoods using loan-level data. However, the regulatory change makes it so treated banks are no longer required to complete the loan reporting requirements associated with the CRA, preventing this detailed analysis.

Instead, I first evaluate whether there were any changes in lending to small businesses and farms for treated banks. While this does not directly capture changes to the distribution of loans based on neighborhood income, it does evaluate a bank's overall lending to their communities. I find that banks affected by the regulatory change to the CRA increase their small business lending.

Even though I find that banks increase their small business lending following the regulatory change, it still could be the case that after the removal of the reporting requirements, banks shift where they are originating small business loans. This argument follows from the same logic as the possibility of banks shifting their lending away from small businesses. While I am unable to use the full sample of loan level data, I use a specific sub-sample of small business lending as a proxy for banks' overall small business lending: Small Business Administration (SBA) guaranteed loans. SBA guaranteed loans are loans for up to $5 million that have a portion guaranteed by the SBA.[4] The guarantee by the SBA decreases the risk exposure to the bank, which makes this portion of a bank's small business lending different from standard small business lending. Importantly, the SBA publishes loan-level data on guaranteed loans, including the location of the loan. This allows me to determine if affected banks changed their lending, especially in lower-income areas. I find no evidence that treated banks shift their small business lending away from low-income areas.

> **Hypothesis 3:** The CRA regulatory change does not negatively impact the quality of bank loans.

One of the key motivations driving the changes to the CRA, according to regulators, was to relieve smaller banks of an undue regulatory burden. I show that the reduction in regulatory oversight associated with the 2005 CRA regulatory change leads to an increase in bank lending for treated banks, suggesting a reduction in regulatory burden. However, a concern could be that banks are increasing their lending by relaxing their lending standards,

[4] For the standard 7(a) loans up to $150,000, the SBA will guarantee up to 85%. For loans over this, the guarantee has a cap of 75%.

i.e. making riskier loans, due to a decrease in regulatory oversight. If a decrease in the quality of the loans occurs following the increase in lending, this would suggest a moderating impact on the effectiveness of the regulatory change. To investigate this, I look at lending quality using both the balance sheet and income statement. A decrease in lending quality would likely be associated with an increase in past-due loans and charge-offs. I find no evidence that following the regulatory change, treated banks have an increase in charge-offs or past-due loans. These findings suggest that the decrease in regulatory oversight associated with the 2005 CRA regulatory change is not associated with a decrease in the lending quality of the bank.

The rest of my paper is organized as follows. Section 2 discusses the related literature. Section 3 provides institutional background on the CRA and FDICIA. Section 4 provides the data and methodology used in my analysis. Section 5 presents the empirical results. Section 6 concludes.

## 2.2 Literature Review

My research contributes to three strands of literature. I first contribute to the growing literature on liquidity creation. The liquidity creation measure developed in Berger and Bouwman, 2009 has led to numerous studies on liquidity creation. The majority of the research has focused on the interaction of liquidity creation and regulatory capital (eg. Berger and Bouwman, 2009; Casu et al., 2019; Fungáčová et al., 2017; Horváth et al., 2014; Tran et al., 2016; Zheng et al., 2019). Investigations into regulatory intervention and competition created by regulatory action suggest that these events decrease liquidity creation (Berger et al., 2016; Jiang et al., 2019). Numerous papers find that bank size impacts liquidity creation (Berger & Bouwman, 2009; Berger et al., 2017; Dang & Dang, 2021). Bank management ability and governance have been found to have a positive relation with liquidity creation (Andreou et al., 2016; Díaz & Huang, 2017). Research has also explored areas such as systematic risk (Davydov et al., 2021), prediction of recessions (Chatterjee, 2018), profitability (Duan & Niu, 2020), and merger activity (Pana et al., 2010). I contribute to this literature by analyzing the impact of compliance with size-based regulatory thresholds on liquidity creation. To my knowledge, the only other paper to do so is Nguyen et al., 2020. They examine liquidity creation around the bank stress test threshold. I contribute by looking at the impact of regulatory thresholds for community banks. In addition, stress testing is inherently linked to bank capital, so I add to the literature by considering regulatory size thresholds that are not directly linked to bank capital.

I also contribute to the literature that evaluates bank actions around size-based regulatory thresholds. Several bank regulations have size-based threshold requirements, such as the CRA, FDICIA, Home Mortgage Disclosure Act (HMDA), and Dodd-Frank. Many studies focus on Dodd-Frank and the related stress test thresholds. Compliance with Dodd-Frank requirements is linked with decreases in profitability in the short run for all banks but has a positive link in the long term (Liu, 2019). Some research finds that compliance with Dodd-Frank regulations decreases risk (Clark et al., 2020; Jiang, 2020). However, there is also some evidence that Dodd-Frank did not decrease systematic risk (Q. Huang, 2018). Banks around the threshold for Dodd-Frank compliance appear to manipulate their growth using deposits and acquisition

(Ballew et al., 2022; Bindal et al., 2020; Bouwman et al., 2018). Research also shows that stress test compliance reduces moral hazard and reduces credit supply prior to the stress test (Acharya et al., 2018; García & Steele, 2022). In addition, closely related to my analysis, Nguyen et al., 2020 finds that regulatory stress tests hurt liquidity creation.

When considering smaller banks, the regulatory changes of the FDICIA and CRA are ideal to examine as they capture a larger number of banks than those affected by the Dodd-Frank thresholds. To the best of my knowledge, the previous literature on smaller banks has focused mainly on the FDICIA regulatory size threshold. Altamuro and Beatty, 2010 find that compliance with the FDICIA increases the validity of loan loss provisions, earnings persistence, and cash flow predictability. Jin, Kanagaretnam, Lobo, and Mathieu, 2013 finds that compliance with the FDICIA is associated with decreased risks and lowered probability of bank failure during the financial crisis. Jin, Kanagaretnam, and Lobo, 2013 finds that after the 2005 regulatory change, banks above the threshold have increased risk and probability of failure. Gopalan et al., 2020 finds that reliance on external verification of internal controls is an imperfect substitute by evaluating changes to CAMELS ratings surrounding the 2005 regulatory change. I add to this literature in two ways. First, I evaluate the impact that both the CRA and FDICIA regulatory changes have on bank liquidity creation. Secondly, I consider the importance of separating the CRA and FDICIA regulatory changes. As both regulations change in 2005, my analysis considers the impact of ignoring the possible contaminating samples, and I show that each event must be considered separately. To my knowledge, this is an area that the literature has previously not addressed.

Finally, I contribute to the literature on regulatory compliance for banks. Regulations often have material costs of compliance. Prior studies have evaluated both the costs and benefits associated with compliance (Alvero et al., 2022; Posner & Weyl, 2013). Cyree, 2016 finds that different regulations have significantly different regulatory burdens. They find that Dodd-Frank had the largest increase in regulatory burden as compared to the FDICIA and Patriot Act. Much of the bank analysis also considers the differential impact of regulations on smaller banks as compared to large ones. Srivastav and Vallascas, 2019 finds that the increase in the threshold for small holding companies boosted lending without any increased risks. Nicoletti, 2018 finds that smaller banks are at a disadvantage, especially following the financial crisis. I add to this literature by examining the changes in bank liquidity creation related to the CRA and FDICIA compliance. Compliance with these regulations does not directly require banks to change their liquidity. However, the literature has shown that changes to regulatory burden do cause changes in bank behavior. Much of the current research on bank costs of compliance is associated with Dodd-Frank, so I contribute by investigating regulatory compliance costs for smaller banks. In addition, I contribute by considering changes to bank behavior following an expected reduction in regulatory burden. Overall, the literature focuses on increases in bank regulatory burden as opposed to the reduction of said burden.

## 2.3 Institutional Background

### 2.3.1 Community Reinvestment Act

The Community Reinvestment Act (CRA) was enacted in 1977 to prevent what is known as redlining. Redlining is the event where certain areas are deemed to be undesirable to lend to, usually due to the race or ethnicity of the people living in the area. Many believe that the term redlining stems from the set of maps created by the Home Owners' Loan Corporation (HOLC) in the late 1930s and early 1940s. The maps rated areas in terms of loan riskiness, and the least desirable areas were colored red.[5] Specifically, the CRA requires that bank supervisory agencies (Federal Reserve, OCC, FDIC, and OTS) evaluate banks to determine if they are "helping to meet the credit needs of its entire community, including low- and moderate-income neighborhoods, consistent with the safe and sound operation of the bank" (12 C.F.R §228.11.b). At the initiation of the act, due to the examination process, all banks were required to maintain loan-level records documenting their compliance.

Due to complaints about the unclear and inconsistently applied policies related to the CRA as well as the excessive paperwork associated with examinations, in 1995, the CRA was revised (60 FR 22190). This update helped to clarify the requirements of the CRA. In addition, the revision implements the concept of "small institutions." As of 1995, banks below $250 million in assets were classified as small institutions[6]. The bank was considered a small institution if, as of December 31 in the previous two years, the bank had assets below

[5]To see the maps and obtain further detail, please see https://dsl.richmond.edu/panorama/redlining

[6]Banks with assets below the threshold that are owned by holding companies with less than $1 billion in assets also were classified as small institutions.

$250 million. Banks classified as small institutions were exempt from the data reporting requirements as described in 12 C.F.R §228.42.

Effective September 1, 2005, further updates were made to the CRA thresholds. This regulatory change created a new category of banks denoted as intermediate small banks (70 FR 44256). A bank is classified as a small bank unless it meets the criteria for an intermediate small bank or a large bank. At the time of the change, a bank was classified as an intermediate small bank when the bank had assets of at least $250 million as of December 31 in both prior years but had assets less than $1 billion in one of the two years. A large bank had assets of at least $1 billion in both of the prior years as of December 31[7]. The thresholds for large and intermediate small banks are adjusted annually for inflation. Small and intermediate small banks do not have to collect and report loan-level data and are subject to less stringent examination procedures.

### 2.3.2 FDICIA

The FDIC Improvement Act (FDICIA) was enacted in 1991 to address numerous concerns in the banking industry. The main focus of the regulation was on ensuring that banks were appropriately monitored for capital adequacy and provided actions to take concerning troubled banks. This led to two major parts of the regulation: the concepts of prompt corrective action and least-cost resolution (12 C.F.R §324 and 12 C.F.R §360). Prompt corrective action allows the FDIC to take progressively more stringent actions against FDIC-insured financial institutions if the institutions are deemed to be undercapitalized. Banks are assigned into five

[7]Holding company size no longer was taken into consideration for the CRA threshold purposes.

categories ranging from well-capitalized to critically undercapitalized. In situations where an institution is determined to be in the lowest three categories, they are subject to action by their supervisory agency to return them to at least an "adequately capitalized" rating. The least-cost resolution says that, in the event of a bank failure, the FDIC should take the solution that is determined to pose the lowest cost to taxpayers. Notably, the least-cost resolution included an exemption for "insured depository institution would have serious adverse effects on economic conditions or financial stability", which became better known as the concept of too big to fail (Richards, 2013).

As another monitoring mechanism, the FDICIA established new regulations requiring independent annual audits (12 C.F.R §363). Initially, the exemption from this part of the regulation was set at $150 million in assets but was increased to $500 million, prior to the act's implementation in 1993 (58 FR 31332). For institutions that have $500 million or more in assets as of the beginning of their fiscal year, the following requirements apply:

1. Audited Financial Statements[8]
2. Independent public accountant's report on audited financial statements
3. A report signed by the chief executive officer and the chief accounting or financial officer including:
    (a) A statement of management's responsibilities for:
        i. Preparing financial statements

[8]The audited financial statements can be those of the top-tier or any mid-tier holding company. As of June 15, 2010, the financial institution must make up 75% of the consolidated assets of the audited holding company. To qualify for this, the services of the holding company must be comparable. In addition, the bank itself either must have assets below $5 billion or a CAMELS rating of 1 or 2.

   ii. Establishing and maintaining an adequate internal control structure and procedures for financial reporting

   iii. Complying with the laws and regulations relating to safety and soundness which are designated by the FDIC or the appropriate Federal banking agency

  (b) An assessment of

   i. The effectiveness of such internal control structure and procedures

   ii. The institution's compliance with the laws and regulations relating to safety and soundness

4. The independent public accountant's attestation report concerning the effectiveness of the institution's internal control structure over financial reporting

5. Establish an independent audit committee made up entirely of outside directors

Effective December 31, 2005, these regulations were updated so that items 3(b)i and 4 were no longer required for banks with assets below $1 billion[9] (70 FR 71226). In addition, the requirements surrounding an independent audit committee were also updated. For banks with assets between $500 million and $1 billion, the audit committee is required to have a majority of outside directors instead of the entire committee being outside directors. This update was driven by the regulatory burden that the FDIC felt smaller institutions faced in compliance with the FDICIA.

The regulations were further updated in 2009 to clarify specific requirements pertaining to 3(b)i and 4 (74 FR 35745). This clarification contained four parts that apply to both manage-

[9] The exception to this being institutions that are publicly traded or are subsidiaries to publicly traded companies required to comply with Sarbanes-Oxley Act Section 404.

ment and the independent public accountants[10]. The first item was that the internal control framework used needed to be explicitly identified. The example given in the regulation is the use of the Committee of Sponsoring Organizations (COSO) of the Treadway Commission, better known as the COSO report. Second, both management and the independent public auditors must state that the "assessment included controls over the preparation of regulatory financial statements in accordance with regulatory reporting instructions and identify the regulatory reporting instructions." Thirdly, both management and the independent public accountants must state their conclusions concerning the effectiveness of the internal controls over financial reporting[11]. Finally, any material weaknesses in the internal controls related to financial reporting that have not been remediated prior to fiscal year end must be disclosed.

Finally, the FDIC implemented a temporary change to the regulation that went into effect October 23, 2020 (85 FR 67433). As a response to COVID-19 and the various cash inflows individuals and businesses received from programs such as the Paycheck Protection Program and stimulus checks, some banks were put above the regulatory thresholds for the FDICIA compliance. According to the FDIC, the payments "may be temporary, but are significant and unpredictable." Because of this, the FDIC allowed institutions to use the lesser of their total assets as of the end of the 2019 calendar year or the beginning of the 2021 fiscal year (85 FR 67427) in determining compliance with the FDICIA. This was specific to the fiscal year 2021 and is no longer listed as part of 12 C.F.R §363.

[10]This also provided more guidance on banks with assets over $3 billion in terms of their audit committee. At this threshold, the audit committee must have members with banking experience, have access to their own outside legal counsel, and not include any large customers of the bank.

[11]Any material weaknesses identified but not remediated prior to the fiscal year-end would require a conclusion that the internal controls were ineffective

## 2.4 Data and Methodology

### 2.4.1 Identification

2005 saw two regulatory changes affecting community banks surrounding size-based thresholds. For the CRA, the regulatory change affected banks above $250 million in assets, and for the FDICIA, banks above $500 million in assets were affected. Often with regulatory changes, such as Dodd-Frank, the treated group are subject to additional compliance requirements. However, the changes to the CRA and FDICIA in 2005, represent a removal of regulatory compliance requirements. Before the regulatory change, banks with assets above $250 million as of December 31 in the previous two calendar years had to provide loan-level reports on lending for compliance with the CRA. In September 2005, this threshold was raised to $1 billion, thus removing the reporting requirements for banks between $250 million and $1 billion. Similarly, as of the end of 2005, banks with assets above $500 million no longer had to comply with the ICFR requirements in the FDICIA. As such, the following two indicator variables are used to identify the treated groups for the two regulatory size thresholds.

- *CRA*: Banks in year *t* that had assets above $250 million as of December 31 in both year *t-1* and year *t-2*.
- *FDICIA*: Banks in year *t* that had assets above $500 million as of December 31 in year *t-1*[12]

[12] The FDICIA compliance is determined on the fiscal year of a bank. However, less than 4% of banks have fiscal year ends different than December 31st as of 2010. However, identifying the FDICIA-compliant banks by assets in any quarter provides similar results.

I focus my analysis on the impact of the deregulation of the CRA and FDICIA. As such, I am most interested in identifying the treatment effect for the subset of banks required to comply before 2005 that were no longer required to comply following the 2005 regulatory changes. However, the asset size of banks is not constant, but changes over time, which could lead to a bank switching from the control group to the treated group.[13] As such, *CRA* (*FDICIA*) are limited to observations where a firm remained treated/untreated for the entire sample period. This is similar to the setup used in Gopalan et al., 2020, however, their treatment group indicator only is based on a bank's compliance requirements as of the end of 2005. I add an extra layer here that does not allow banks to switch between the treatment and control group throughout my sample. My results, however, are robust to allowing banks to switch treatment groups.

At $1 billion in assets, the regulatory compliance components removed for both the CRA and FDICIA come back into effect. In addition, according to the FDIC, banks below $1 billion in assets are identified as community banks, which is the focus of my analysis. As such, I exclude observations where the bank has assets over $1 billion.[14] Both regulatory changes were effective by the end of 2005, so *Post* is an indicator variable for years after 2005. It should be noted that the initial announcement for both regulatory changes occurred earlier in 2005. It is possible that banks altered their liquidity creation earlier in 2005 in anticipation of the regulation change being finalized, however, my results are robust excluding 2005 in my sample.

[13] Banks could also switch from the treated to control group, but banks most often grow in size rather than shrink.

[14] In addition, I am unable to separate the two regulation-specific effects for banks above the $1 billion threshold, further supporting the use of a sample with assets below this threshold.

### 2.4.2 Generalized Difference-in-Differences Specification

I use a difference-in-differences design to investigate the impact of changes in regulatory requirements on the liquidity creation of community banks. The use of the 2005 regulatory changes allows me to analyze the liquidity measures for the treated banks in the pre- and post-period. These can then be compared to the control group. The difference-in-differences design allows comparison to a benchmark and enables the identification of the impact of the regulatory change. After assigning banks to the treatment and control groups, I estimate the following generalized regression:

$$LiquidityCreation_{i,t} = \beta_0 + \beta_1 Post * Treated_{i,t} + \delta X_{i,t-1} + \lambda_i + \gamma_{st} + \epsilon_{i,t} \tag{2.1}$$

In equation 2.1, $Treated_{i,t}$ is either *CRA* or *FDICIA* as previously defined. The primary variable of interest is $Treated_{i,t} * Post$. *Post* is an indicator equal to one for quarters in years after 2005. $X_{i,t-1}$ includes a number of lagged control variables. I control for bank size (natural logarithm of total assets), total risk-based capital ratio (total capital/risk-weighted assets), ROA (net income/total assets), and non-performing loans (total non-performing loans/total loans). All income variables are annualized by using the sum of the current quarter and the prior three quarters.[15] I also include both bank and state-by-year-quarter fixed effects. The bank (quarter) fixed effects cause $Treated_{i,t}$ ($Post$) to be absorbed. The bank-fixed effects capture bank-specific but time-invariant characteristics. The state/time

[15] Call reports report income on a cumulative basis throughout the year. I first identify the quarter-specific number and then annualize.

fixed effects capture possible shocks that vary over time and are region specific which may bias my results.[16]

### 2.4.3 Regulatory Overlap

One of the main contributions of my paper is my analysis of the consequence of treating the FDICIA and CRA regulatory changes as independent events. Both regulatory changes occurred in 2005 and impacted banks of a similar size. Generally, with an event such as a regulatory change, there are two groups: the treated and the control group. This allows for the use of a differences-in-differences (DiD) regression. However, concerning the 2005 regulatory changes, there are actually three groups. Figure 2.1 shows this breakdown. The figure shows the split using the size-based thresholds for the CRA and FDICIA. As shown, this leads to three groups. Group 1 contains all banks with assets below $250 million in assets. These banks are not required to comply with the CRA reporting or FDICIA audit and ICFR requirements before or after 2005. Group 2 contains banks above $250 million and below $500 million in assets. These banks are not required to comply with the FDICIA audit and ICFR requirements before or after 2005. However, group two banks were required to comply with the CRA reporting requirements before 2005. Group 3 consists of banks above $500 million and below $1 billion in assets. These banks were required to comply with the CRA reporting and FDICIA audit and ICFR requirements prior to 2005, but are only required to comply with the FDICIA audit requirement following the 2005 regulatory changes.

[16] My results are robust to variations in fixed effects.

If each regulatory change were treated as an independent event, a treatment and control group would be assigned using the size-based threshold. For the FDICIA regulatory change, this would mean that all banks above \$500 million in assets would be in the treated group (Group 3) and all banks below \$500 million would be in the control group (Groups 1 and 2). Similarly, for the CRA regulatory change, the treated group would consist of all banks with assets above \$250 million in assets (Groups 2 and 3), and all banks below \$250 million would be the control group (Group 1). The estimates of the DiD regressions then could be expressed as[17]:

$$\begin{aligned} DiD_{FDICIA} &= Diff[\text{Group 3}] - Diff[\text{Groups 1 and 2}] \\ &= \Delta\text{Group 3} - \frac{\alpha_1\Delta\text{Group 1} + \alpha_2\Delta\text{Group 2}}{\alpha_1+\alpha_2} \\ &= \Delta CRA + \Delta FDICIA - \frac{\alpha_2\Delta CRA}{\alpha_1+\alpha_2} \end{aligned} \tag{2.2}$$

$$\begin{aligned} DiD_{CRA} &= Diff[\text{Group 2 and 3}] - Diff[\text{Group 1}] \\ &= \frac{\alpha_2\Delta\text{Group 2} + \alpha_3\Delta\text{Group 3}}{\alpha_2+\alpha_3} - \Delta\text{Group 1} \\ &= \frac{\alpha_2\Delta CRA + \alpha_3(\Delta CRA + \Delta FDICIA)}{\alpha_2+\alpha_3} \end{aligned} \tag{2.3}$$

$\alpha_j$ represents the distribution between the two groups. The key takeaway from the two equations is that neither cleanly identifies the FDICIA-specific and CRA-specific effects. Failure to properly separate the effects confounds the findings. Instead, I propose using sub-samples of the data to alleviate the issues with the overlap. I start with the CRA analysis as it is more straightforward. Instead of setting the treated group as all banks with assets

[17]This setup/explanation follows the logic described in Faulkender and Petersen, 2012.

above \$250 million, I limit my treated group to banks with assets above \$250 million but below \$500 million (Group 2). The control group remains the same (Group 1). Now the DiD regression is:

$$\begin{aligned} DiD_{CRA} &= Diff[\text{Group 2}] - Diff[\text{Group 1}] \\ &= \Delta\text{Group 2} - \Delta\text{Group 1} \\ &= \Delta CRA \end{aligned} \tag{2.4}$$

This identifies the CRA-specific effects without any possible contamination from the FDICIA regulatory change. To identify the FDICIA-specific effects, Group 3 must remain as the treated group. However, if Group 3 is the treated group and Group 1 is the control group, as in the CRA analysis, the CRA-specific effects are still not accounted for. Instead, in the FDICIA analysis, the treated group is Group 3, and the control group is Group 2. The equation then becomes:

$$\begin{aligned} DiD_{FDICIA} &= Diff[\text{Group 3}] - Diff[\text{Group 2}] \\ &= \Delta\text{Group 3} - \Delta\text{Group 2} \\ &= \Delta\text{CRA} + \Delta FDICIA - \Delta CRA \\ &= \Delta FDICIA \end{aligned} \tag{2.5}$$

### 2.4.4 Data

I use the Consolidated Reports of Condition and Income (Call reports) for the period of 2000-2010 as the primary source of data in my analysis. Call reports are quarterly reports filed by banks that include a balance sheet, income statement, and a number of supporting schedules related to the activities of the bank such as lending and deposit activity. Banks of all sizes must complete the report on a quarterly basis, though the amount of information provided varies with the size, activities, and complexity of the bank. Following Gopalan et al., 2020, I require that in a given quarter $t$, the bank not be missing total assets in the current quarter or in the prior three quarters. The final sample contains 224,581 bank-quarter observations. In my preferred sub-samples, 9.48% of the 213,126 observations in the CRA analysis make up the treated sample. 36.19% of the FDICIA sample (31,656 observations) are treated. All variables in my analysis are winsorized at the 1% and 99% levels.

I follow the methodology of Berger and Bouwman, 2009 to construct liquidity creation measures. Table 2.2 shows the specific variables used and the creation of the measures. Berger and Bouwman, 2009 develops two liquidity creation measures–"cat fat" and "cat no fat"– which capture liquidity creation from both on- and off-balance sheet items and on-balance sheet only items, respectively.[18] I use the notation in Nguyen et al., 2020 to denote total liquidity creation as *LC(Total)* and on-balance sheet liquidity creation as *LC(On)*. The

[18]The paper has two further measures that include maturity-related analysis.

asset, liability, and off-balance sheet components are denoted as *LC(Asset)*, *LC(Liability)*, and *LC(Off)*, respectively. All measures are scaled by Gross Total Assets (GTA).[19][20]

### Small Business Lending Outcomes

The overarching goal of the CRA is to ensure that banks address the credit needs of the entire community, including lower-income areas. If the regulatory change leads to bank behavior that is contrary to this goal, it could evidence unintended negative consequences. Ideally, this would be tested using loan-level data to identify any changes to the lending habits of treated banks following the regulatory change. However, the CRA regulatory change is such that treated banks are no longer required to report loans for the CRA purposes. As such, I cannot use CRA files with loan-level data. Instead, I use the annual small business and farm lending data reported in the call report which are measures of overall lending to the community (i.e small businesses). This is reported on an annual basis in June.[21] The data is found in schedule RC-C part II and is divided into information about small business lending and loans to small farms. I analyze the number and size of the loans outstanding for each category as well as the combination of the two. Small business loans are identified as loans with origination amounts of $1 million or less and small farm loans are loans with origination amounts of $500,000 or less in their appropriate categories.

[19] GTA is defined as total assets plus allowance for loan and lease losses (ALLL) and allocated transfer risk reserve.

[20] After winsorization, my measures correlate at levels above 0.99 with those provided at https://sites.google.com/site/hughhkimswebsite/home.

[21] In 2010, banks began reporting this information on a quarterly basis.

**SBA Loans**

While I am unable to use CRA loan-level reports to test changes in lending decisions, I can analyze a small subset of small business loans at the loan level: SBA-guaranteed loans. Specifically, I utilize the loan-level data from the SBA 7(a) loan program. According to the SBA website, these are the most common loans guaranteed by the SBA. The loans are originated by the bank and up to 85% of the loan value is guaranteed by the SBA.[22] The data provides the location of the project, the amount of the loan, and the bank originating the loan. Specifically, I want to analyze any changes in the lending behavior of treated banks for lower-income and higher-income areas. To do so, I also utilize the Federal Financial Institutions Examination Council (FFIEC) census and demographic data.

The census and demographic data from the FFIEC includes income level classifications at the census tract level as defined by the CRA. The income classifications identify if a census tract is considered to be CRA-eligible or not. CRA-eligible tracts are census tracts identified as distressed or underserved. Eligible tracts are the locations for which banks need to provide evidence providing for the credit needs of borrowers. Low and moderate-income tracts are considered to be CRA-eligible.[23] I aggregate the CRA eligibility from the tract to the zip-code level by taking the weighted average, by area, of the income level classifications. I then match these to the zip code in which the SBA loan project is located. I finally aggregate the data

[22]For the standard 7(a) loans up to $150,000, the SBA will guarantee up to 85%. For loans over this, the guarantee has a cap of 75%.

[23]Low-income tracts are those with a median income below 50% relative to either the median family income of the metropolitan statistical area (MA)/statewide non-MA (MFI). Moderate income tracts are those with a median income between 50% and 80% of MFI. Middle-income tracts are those with a median income between 80% and 120% of MFI. Upper-income tracts are those with a median income above 120% of MFI.

to the bank-year-CRA eligibility level to obtain the number and total size of loans originated by the bank to CRA-eligible and CRA ineligible zip codes.

## 2.5 Summary Statistics and Results

### 2.5.1 Summary Statistics

Table 2.3 shows the summary statistics for my sample.[24] The banks on average are relatively small, with average assets of $156.44 million. On average, the combined on- and off-balance sheet liquidity creation in my sample is 27.7% of GTA. This is similar to that described for small banks in Berger and Bouwman, 2009. Off-balance sheet liquidity creation is 5.0% of GTA and is almost entirely composed of illiquid guarantees. This shows that most liquidity creation for community banks is driven mostly by on-balance sheet activity, consistent with Berger and Bouwman, 2009. I however still analyze the impact of the 2005 regulatory changes on total liquidity creation, LC(Total).[25] Much of the liquidity creation comes from liquid liabilities and illiquid assets as these are 44.1% and 44.6% of GTA, respectively.

Overall, the banks in my sample appear to be healthy. The banks are also on average profitable with an annual ROA of 0.9%. The banks are sufficiently capitalized, with their risk-based capital ratio being well over the required 8%. Less than 2% of loans are considered

[24]Summary statistics for the CRA and FDICIA specific groups are provided in Appendix A.2. Other than obvious differences in the total assets, the samples appear relatively similar.

[25]My results are similar if I analyze LC(On) instead)

to be non-performing on average and this holds even in the 75th percentile. Banks in the sample have annualized charge-offs of less than 1% of total loans.

Small business and farm lending makes up, on average, $39.67 million dollars of the banks' overall outstanding loan portfolios. This is mostly concentrated in small business lending, with only 21.6% of this concentrated in small farm lending. The overall volume of small business and farm lending is distributed between, on average, 473 loans.

### 2.5.2 Parallel Trends

One of the key criteria that allow any conclusions to be drawn using the difference-in-differences methodology is the assumption of parallel trends. This is the requirement that absent of treatment, the treated banks would have followed the same trend as the control banks. While this is untestable, I conduct several analyses to provide support for this assumption. I first plot the univariate average for liquidity creation over the sample period. Figures 2.3 and 2.4 show the results. These figures plot the univariate averages for the treated and control banks for each regulatory change. The graphical evidence suggests that the slopes between the two groups are mostly parallel prior to the regulatory change. However, to provide more robust support for the parallel trends, I follow the methodology used in Bindal et al., 2020[26]. To do so, I estimate the following equation:

$$LiquidityCreation_{i,t} = \beta_0 + \beta_1 Treated_{i,t} * TimeTrend + \lambda_i + \gamma_{st} + \epsilon_{i,t} \tag{2.6}$$

[26] I maintain the fixed effects at the state by year-quarter level, so my regression is slightly altered from their setup.

$LiquidityCreation_{i,t}$ and $Treated_{i,t}$ are the same as in 2.1. *TimeTrend* is the current quarter of the observation less the quarter ending December 31, 1999. The sample for the analysis includes the period before the regulatory change implementation, so 2000-2005. The important coefficient in this regression is the interaction term: $Treated_{i,t} * TimeTrend$. If this is insignificant, it supports the required assumption that there is no significant difference in the liquidity creation trend between the treated and control group prior to treatment. I evaluate this for the baseline full sample as well as the non-regulatory overlap sub-samples. Table 2.4 shows these results. The interaction term is insignificantly different from zero for both regulatory changes. This holds for both the full sample estimation, in columns 1 and 3 and the preferred sub-sample estimation in columns 2 and 4, supporting the parallel trends assumption.

As a further robustness test, I estimate equation 2.1. However, I split *Post* into separate indicator variables for each year ($\delta_t$). The coefficients before the regulatory changes should exhibit no trend and be insignificantly different from zero. I treat 2004 as the excluded year in the following equation:

$$LiquidityCreation_{i,t} = \beta_0 + \sum_{t=2000}^{2010} \beta_t Treated_{i,t} * \delta_t + \lambda_i + \gamma_{st} + \epsilon_{i,t} \tag{2.7}$$

Figures 2.5 and 2.6 show the analysis for each regulatory change. In each, the left figure shows the analysis when each regulation is considered independently. The right figure shows the analysis using my preferred sub-samples that address the regulatory overlap. In all situations, the coefficients are zero and show no notable trends before 2006, further supporting the parallel trends assumption.[27]

### 2.5.3 Liquidity Creation and Regulatory Overlap

Table 2.5 presents the results of the investigation into the relationship between the change in the regulatory size thresholds for the CRA and FDICIA. This table analyzes my baseline specification where each regulation is treated independently. I then compare these results to analysis using my preferred specification to address the regulatory overlap. Columns 1 and 2 and 5 and 6 evaluate the regulatory changes related to the CRA and FDICIA when treating them as independent events, respectively. Under hypothesis 1, I expect that the coefficients on the interaction term should be positive. The coefficients in the regression support this hypothesis. Banks that are affected by the removal of ICFR requirements or decreased reporting requirements associated with the CRA appear to increase their liquidity creation following both regulatory changes. Economically, these results would suggest that treated banks increase their liquidity creation relative to the control group by 5.78% relative to the unconditional mean of *LC(Total)* following the CRA regulatory change. Similarly, these results suggest that banks increase their liquidity creation following the removal of ICFR requirements by 6.86% relative to the unconditional mean.

As previously discussed in Section 2.4.3, the overlap between the two samples makes it so the CRA-specific effects and FDICIA-specific effects cannot be cleanly identified if each regulatory change is treated independently. Without making adjustments, the treated group

[27] In untabulated results, using quarterly indicators yields the same results.

for the CRA regulation consists of banks both affected and unaffected by the FCICIA ICFR removal. Similarly, the control group for the FDICIA regulation consists of parts of both the CRA-treated and control groups. Both of these situations could lead to confounding results when evaluating the impact of regulatory change on liquidity creation. I correct for the overlap in the CRA analysis by excluding the FDICIA-treated banks. This excludes any FDICIA-specific effects from my analysis. For the FDICIA results, I cannot exclude the treated group from the CRA regulation as all the treated banks under the FDICIA are also treated under the CRA regulatory change. Instead, I exclude any banks in the control sample that is part of the CRA control group. This removes any CRA-specific effects from my analysis of the FDICIA regulation.

Columns 3 and 4 and 7 and 8 of Table 2.5 show the results of my non-overlapping samples for the CRA and FDICIA regulations, respectively. Columns 3 and 4 show that the CRA regulation change leads to a positive and significant increase in liquidity creation with and without controls. The results also suggest that relative to the unconditional mean of *LC(Total)*, banks that no longer are required to report loans for the CRA compliance increase their liquidity creation by 5.15% relative to the unconditional mean of *LC(Total)* in the preferred CRA sample.[28] This shows that failing to consider the possibility of the FDICIA-specific effects would result in a slight overestimation of the magnitude of the CRA-specific effects.

[28]The mean of LC(Total) is 27.2% of GTA for the CRA sub-sample. See Appendix A.2 for the full list of summary statistics.

When evaluating the impact of the FDICIA regulatory change after excluding the CRA-specific effects, columns 7 and 8 show that the interaction term is insignificantly different from zero, confirming the necessity of properly addressing the regulatory overlap. Without properly correcting for the contaminating effects, the increase in liquidity creation driven by the regulatory changes would be overstated and incorrectly attributed to both the CRA and FDICIA regulatory changes, which further supports the importance of correctly conducting this analysis. Due to this finding, all further analysis will use the CRA sub-sample.[29]

### 2.5.4 The Drivers of Liquidity Creation

As previously discussed, the basic formula that describes how banks create liquidity is they use liquid liabilities to finance illiquid assets. However, in terms of the total liquidity creation of the bank, it is important to consider all the components, as there is still an interplay between illiquid liabilities, liquid assets, and off-balance sheet liquidity creation. By considering the breakdown, I provide a more granular analysis of my previous findings in Section 2.5.3. This analysis shows the exact mechanisms driving this liquidity creation.

Table 2.6 analyzes the liquid and illiquid components of Berger and Bouwman, 2009's liquidity creation measure. I exclude liquid derivatives and liquid guarantees as these are almost never utilized by community banks. I found previously in Table 2.5 that banks have positive and significant liquidity creation following the CRA regulatory change. The results show that the increase in liquidity is being driven by, as expected, an increase in liquid liabilities that are being converted into illiquid assets as both coefficients are positive and

[29] Appendix A.2 provides versions of key tables discussed in the paper for the FDICIA sub-sample.

significant. I also find a statistically significant increase in illiquid guarantees, which is consistent with an increase in liquidity creation. I finally find that there is a decrease in liquid assets. A decrease in liquid assets also is a signal of liquidity creation per the Berger and Bouwman, 2009 methodology.

### 2.5.5 Small Business Lending Outcomes

I have shown that following the 2005 CRA regulatory changes, treated banks increase their liquidity creation. In Table 2.6, I further show that banks increase their liquidity creation by increasing both their liquid liabilities and illiquid assets. Illiquid assets are, in part, made up of commercial real estate (CRE), commercial and industrial loans, and loans to finance agricultural production. Loans below certain sizes within these three categories are considered to be small business and small farm loans. According to the Federal Reserve, the purpose of the CRA is "to encourage financial institutions to help meet the credit needs of the communities in which they do business[30]." Even though I have shown overall that the 2005 regulatory change in the CRA increased lending for affected banks, it is important to look at the types of loans that are important to the goals of the CRA. After the 2005 regulation change, affected banks have reduced transparency as they no longer have to provide detailed loan-level reports. This increase in opacity, if followed by a decrease in small business/farm lending, would suggest that the increase in lending did not translate to small business/farm lending, which would be contrary to the goals of the CRA. It could be the case that banks view small business/small farm lending as riskier and use the increased opacity to shift away

[30]https://www.federalreserve.gov/consumerscommunities/cra_about.htm

from these risky loans to something they deem to be safer. I conduct this analysis using the same generalized DiD equation previously discussed in Equation 2.1.

Table 2.7 analyzes the impact of the 2005 CRA regulatory change on small business and farm lending. I consider the aggregate of the sub-categories small business and small farm lending. Appendix A.2 provides the breakdown of each category within small business and small farm lending. Overall, I find that affected banks significantly increase their total small business/farm lending, on average by $4.449 million, which is an increase of 14.41% relative to the mean of the CRA-specific sample ($30.88 million). This is driven by increases in small business lending. The coefficient on the small farm lending is also positive, but it is insignificant. The results show that the increase in liquidity creation is not done solely by non-small business/farm loans, suggesting that the regulatory change does not have outcomes contrary to the purpose of the CRA.

**SBA Lending**

The previous analysis considers changes to banks' overall small business lending. My findings suggest that the goal of the CRA to encourage banks to lend to the community, as a whole, is not negatively affected by the regulatory change. However, a second critical component of the CRA is making sure that the needs of lower-income areas in the community have their credit needs addressed. Previous literature uses variations in the census tract income eligibility to show that overall, the CRA increases small business lending (Bostic & Lee, 2017; Ding et al., 2020; Kim et al., 2021). My analysis is different from this as I do not use variations in census tracts to evaluate the CRA effectiveness.[31] Instead, I consider whether the decrease in regulatory requirements negatively impacts the effectiveness of the goals of the CRA.

Using the subset of small business lending loan data from the SBA loan files, I identify loans associated with projects in CRA-eligible and ineligible zip codes. This allows me to test if there are changes to banks' lending behavior in low-income zip codes. For the analysis, I first aggregate the loan-level data to the bank-year- zip code level. I then divide the data into CRA-eligible and non-eligible zip codes. As such, each bank can have at most two observations per year. I conduct my analysis on each sample (CRA eligible and ineligible) separately. In this setup, I can identify any changes in the treated banks' lending behavior for the CRA-eligible zip codes, which are the main interest in my analysis.

I analyze the number of loans the bank originates, the total volume of SBA loans originated, the total number of jobs estimated to be supported by the loans, and the percentage of loans that were charged off. Table 2.8 shows the results of the analysis. Columns 1-4 represent the loans in CRA-eligible zip codes while columns 5-8 represent those in the remaining zip codes. While the results for the non-CRA-eligible zip codes are interesting, the main focus is on the CRA-eligible (underserved) zip codes. I find no evidence of significant changes in SBA lending in underserved counties. None of the coefficients for my interaction term are significantly different from zero. This suggests that small business lending to low-income areas did not decrease, which would be the major concern, following the 2005 regulatory change. The key

[31] CRA eligibility is determined at the census tract level.

finding here is that it does not appear that the decrease in required regulatory reporting lead to behavior by affected banks that was contrary to the goals of the CRA.

Interestingly, when we analyze the loans to the remaining zip codes, I find that the number of jobs supported by SBA loans increases significantly for banks affected by the CRA regulatory change. I also show in Table A.7 in Appendix A.2 that the number of jobs supported and the number of loans originated are statistically significantly larger for non-CRA-eligible zip codes relative to CRA-eligible ones. These results are not necessarily contrary to the goals of regulators in reducing the CRA reporting requirements for eligible banks, as the concern would be if there was evidence of a decrease in lending to low-income areas. However, these results do point to evidence that banks lend more, in general, to higher-income areas. Banks likely perceive loans to CRA-eligible areas to be more costly, so while the removal of oversight from the CRA reporting does not appear to reduce lending to low-income areas, it does not appear that banks increase lending equally across CRA-eligible and CRA-non-eligible zip codes.

### 2.5.6 Lending Quality

Throughout my analysis, I have shown that, following the CRA regulatory change, treated banks increase their lending overall as well as for their small business lending portfolios. This supports the conclusion that the decrease in reporting requirements following the CRA regulatory change led to a reduction in regulatory burden for the bank, which was the goal of the regulators. This reduction in regulatory burden leads to an increase in lending, one of the key purposes of a bank. These results alone would suggest that regulators were successful in achieving a reduction in regulatory burden for impacted banks. However, this accomplishment would be diminished if this increase in lending was associated with banks taking greater risks and accepting lower-quality loans. I consider measures of changes in loan quality from both the income statement and balance sheet to investigate if lending quality appears to change following the CRA regulatory amendment.

In situations where a borrower is 30 days past due on a loan, banks must classify a loan as past due by 30-89 days. Subsequently, if payment has not been made within 90 days, the loan is re-classified as past due by 90+ days. While a loan is classified as past due, it continues to accrue interest, and according to the bank, it is still deemed to be a collectible loan. Loans are transferred to non-accrual status by the bank when at least part of the loan is deemed uncollectible, at which point, the loan stops accruing interest.[32] When a loan is ultimately deemed uncollectible by the bank, the bank removes the loan from the balance sheet by expensing it as a charge-off. If a bank begins to make riskier or lower quality loans, then the number of loans past due, in non-accrual status, and that the bank ultimately has to charge off would also be expected to increase.

Table 2.9 shows the results of changes to past due and non-accrual loans. Columns 2 and 3 show no evidence of an increase in the loans past due by either 30-80 days or 90+ days for treated banks. There is a positive and statistically significant increase in the amount of non-performing loans (column 1) which is driven by the increase in non-accrual loans (column 4) as non-performing loans are the combined total of loans past due by 90+ days and non-accrual loans. If a bank experienced a decline in loan quality, the number of past

[32] Banks are not required to wait until a loan is past due by 90 days to place it in non-accrual status.

due loans would be expected to increase as well as the non-accrual/non-performing loans. Gopalan et al., 2020 provide an in-depth analysis of this pattern and suggest that the increase in non-accrual loans is due to bank discretion as opposed to a decline in loan quality. This is further supported by the results in Table 2.10 which analyzes bank charge-offs. I find no evidence of any change in the amount of charge-offs for treated banks following the 2005 CRA regulatory change. This further supports the finding that bank loan quality does not deteriorate following the decrease in oversight due to the removal of the CRA reporting requirements.

## 2.6 Conclusion

In this study, I evaluate the impact of decreases in regulatory requirements for community banks on liquidity creation. Specifically, I consider the 2005 changes to the FDIC Improvement Act (FDICIA) and the Community Reinvestment Act (CRA) that lessened the regulatory requirements for affected banks. I use a differences-in-differences approach to investigate changes in the liquidity creation for these banks. I show that analyzing the FDICIA and CRA independently leads to confounding results that suggest changes to both regulations lead to an increase in liquidity creation. However, the FDICIA and CRA regulatory changes affect banks of a similar magnitude in the same year, leading to concerns that this analysis does not properly identify the FDICIA- and CRA-specific effects. When I analyze the two regulations without overlap, I find that only the CRA regulatory change leads to increases in bank liquidity creation of 4.20%. Failure to appropriately separate the FDICIA- and CRA-specific effects leads to an overstatement in liquidity creation.

I show that the change in liquidity creation is driven by an increase in liquid liabilities (e.g. deposits) that are then converted into illiquid assets (e.g. loans). As I show that this is driven by the CRA regulatory changes, it is important to consider if this negatively impacts lending to the community, the key purpose of the CRA. As the banks in my sample do not have to report their CRA lending, I use small business and farm lending as a proxy. I find that following the regulatory change, the amount of small business and farm lending increases. In addition, using loan-level data for a subset of the overall small business lending, I find no that affected banks decrease their small business lending in CRA-eligible counties following the CRA regulatory change.

Finally, I investigate whether the increase in lending following the 2005 CRA regulatory change is associated with a change in the bank's credit risk. It could be the case that following the removal of the CRA reporting requirement, banks increase lending due to an increase in riskier loans. If this happened, an increase in past due loans and loan charge-offs would be expected. I however find no evidence of this, suggesting that the increase in bank lending was not driven by banks accepting riskier loans.

Table 2.1: Bank Regulatory Thresholds

This table provides the list of regulatory thresholds for banks with assets under $20 billion (Labonte & Perkins, 2021). The compliance level lists the highest level of consolidation that the regulation applies to. *Bank Holding company directors are required to notify the regulatory agencies prior to dual service. **Thresholds are updated on an annual basis. This is the 2022 number. ***CRA thresholds are updated on an annual basis. This is the 2022 number. Prior to September 2005, the large bank threshold was $250 million for independent banks and $1 billion for bank holding companies but was increased to $1 billion irregardless of holding company status (69 FR 51155).

| Regulation | Threshold | Compliance Level |
|---|---|---|
| Interlocks Act | $50M, $10B | Bank* |
| HMDA** | $50M | Bank |
| BHC Insurance Activities | $50M | Holding Company |
| SEC reporting requirements | $150M | Bank |
| Deposit Insurance Processing | $150M | Holding Company |
| Expedited acquisition activity | $300M, $3B, $7.5B | Holding Company |
| CRA*** | $346M, $1.384B | Bank |
| BHC risk-based capital | $500M | Holding Company |
| FDICIA | $500M, $1B, $3B | Bank |
| Flood Insurance escrow | $1B | Bank |
| TILA** | $2.336B | Bank |
| Collins Amendment | $3B | Holding Company |
| 18 month examination cycle | $3B | Bank |
| Examination Type for BHC | $3B, $10B | Holding Company |
| Stock buyback and redemption | $3B | Holding Company |
| Ineligible to file FFEIC 051 | $5B | Bank |
| CFPB regulator for consumer compliance | $10B | Consolidated |
| Durbin Amendment | $10B | Consolidated |
| Volcker Rule | $10B | Consolidated |
| Portfolio QM | $10B | Bank |
| Mortgage escrow | $10B | Bank |
| CBLR eligibility | $10B | Consolidated |
| Swap margin and capital requirements | $10B | Bank |
| Thrift Charter opt out | $20B | Bank |

Table 2.2: Liquidity Creation and Hoarding Components

This table provides the components used to generate the bank liquidity creation and hoarding measures used in this paper as defined in Berger and Bouwman, 2009. Total bank liquidity creation is defined as LC(Total)=LC(Asset)+LC(Liability)+LC(Off), where LC(Asset)=$\frac{1}{2}$Illiquid Assets-$\frac{1}{2}$Liquid Assets, LC(Liability)=$\frac{1}{2}$Liquid Liabilities-$\frac{1}{2}$Illiquid Liabilities, and LC(Off)=$\frac{1}{2}$Illiquid Guarantees-$\frac{1}{2}$Liquid Guarantees-$\frac{1}{2}$Liquid Derivatives.

| Illiquid Assets | Liquid Assets |
|---|---|
| Commercial real estate loans (CRE) | Cash and due from other institutions |
| Loans to finance agricultural production | All Securities (regardless of maturity) |
| Commercial and industrial loans | Trading assets |
| Other Loans and Lease financing receivables | Fed Funds Sold |
| Other Real Estate Owned (OREO) | |
| Customers' liability on bankers' acceptances | |
| Investment in unconsolidated subisidiaries | |
| Intangible Assets | |
| Premises | |
| Other Assets | |
| **Illiquid Liabilities and Equity** | **Liquid Liabilities and Equity** |
| Bank's liability on bankers' acceptances | Transaction deposits |
| Subordinated Debt | Savings Deposits |
| Other Liabilities | Overnight Fed Funds Purchased |
| Equity | Trading Liabilities |
| **Illiquid Guarantees** | **Liquid guarantees** |
| Unused Commitments | Net participations acquired |
| Net standby letters of credit | |
| Commercial and similar letters of credit | |
| All other off-balance sheet liabilities | |
| **Illiquid derivatives (gross value)** | **Liquid Derivatives (gross value)** |
| | Interest rate derivatives |
| | foreign exchange derivatives |
| | Equity and Commodity derivatives |

Table 2.3: Summary Statistics

This table provides the summary statistics of the variables used in the analysis. All variables, excluding total assets, are winsorized at the 1% and 99% level.

| | N | Mean | SD | P25 | Median | P75 |
|---|---|---|---|---|---|---|
| Total Assets | 224581 | 156.438 | 165.342 | 56.427 | 98.263 | 177.348 |
| Non-Performing Loans | 224568 | 0.013 | 0.017 | 0.002 | 0.007 | 0.017 |
| ROA | 224581 | 0.009 | 0.008 | 0.006 | 0.010 | 0.014 |
| ROE | 224581 | 0.091 | 0.089 | 0.059 | 0.095 | 0.134 |
| Risk Based Cap. Ratio | 224581 | 0.169 | 0.073 | 0.122 | 0.147 | 0.190 |
| Liquidity Creation | | | | | | |
| LC(Total) | 224581 | 0.277 | 0.178 | 0.158 | 0.283 | 0.401 |
| LC(On) | 224581 | 0.227 | 0.156 | 0.126 | 0.236 | 0.337 |
| LC(Off) | 224581 | 0.050 | 0.035 | 0.024 | 0.043 | 0.067 |
| Liquid Assets | 224581 | 0.322 | 0.153 | 0.207 | 0.300 | 0.415 |
| Illiquid Assets | 224581 | 0.446 | 0.171 | 0.321 | 0.444 | 0.569 |
| Liquid Liabilities | 224581 | 0.441 | 0.124 | 0.356 | 0.433 | 0.518 |
| Illiquid Liabilities | 224581 | 0.112 | 0.032 | 0.090 | 0.104 | 0.125 |
| Liquid Derivatives | 224581 | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 |
| Liquid Guarantees | 224581 | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 |
| Illiquid Guarantees | 224581 | 0.100 | 0.069 | 0.049 | 0.085 | 0.135 |
| Small Business and Farm Lending | | | | | | |
| Number Business Loans | 56004 | 320.892 | 296.565 | 129.000 | 230.000 | 397.000 |
| Number Farm Loans | 56004 | 152.421 | 211.971 | 5.000 | 69.000 | 216.000 |
| Number All Loans | 56004 | 473.313 | 399.861 | 211.000 | 359.000 | 602.000 |
| Balance Outstanding Business Loans | 56004 | 27.202 | 28.204 | 8.130 | 17.623 | 35.424 |
| Balance Outstanding Farm Loans | 56004 | 7.508 | 10.548 | 0.371 | 3.352 | 10.493 |
| Balance Outstanding All Loans | 56004 | 34.710 | 30.385 | 14.448 | 25.387 | 44.025 |
| SBA Lending | | | | | | |
| Number Loans | 12414 | 3.570 | 7.729 | 1.000 | 2.000 | 3.000 |
| Loan Size | 12414 | 1.150 | 2.278 | 0.150 | 0.380 | 1.087 |
| Bank Portion | 12414 | 0.278 | 0.552 | 0.024 | 0.090 | 0.262 |
| Jobs Supported | 12414 | 34.074 | 87.264 | 0.000 | 1.000 | 24.000 |
| % Loan Failure | 11661 | 0.166 | 0.299 | 0.000 | 0.000 | 0.227 |
| Loan Quality | | | | | | |
| Past Due Loans 30-89 | 208187 | 0.014 | 0.014 | 0.004 | 0.011 | 0.020 |
| Past Due Loans 90+ | 224568 | 0.004 | 0.006 | 0.000 | 0.001 | 0.004 |
| Non-Accrual Loans | 224568 | 0.010 | 0.015 | 0.001 | 0.004 | 0.012 |
| Charge-Offs | 224568 | 0.004 | 0.007 | 0.001 | 0.002 | 0.005 |
| Recoveries | 224568 | 0.001 | 0.001 | 0.000 | 0.000 | 0.001 |
| Net Charge-Offs | 224568 | 0.003 | 0.007 | 0.000 | 0.001 | 0.004 |

Table 2.4: Parallel Trends Analysis

This table tests the parallel trends assumption for the CRA and FDICIA regulatory change. The sample period is limited to prior to 2005. *CRA* is an indicator variable equal to 1 if the bank had assets above $250 million at the beginning of the current year and previous year, and *FDICIA* is an indicator variable equal to 1 if the bank had assets above $500 million at the beginning of the current year. *Time trend* follows the methodology of Bindal et al., 2020 and is the difference between the current quarter and the quarter ending on December 31, 1999. Columns 1 and 3 uses all observations, and columns 2 and 4 exclude banks that are required to comply with FDICIA and only banks previously required to comply with CRA, respectively. Bank and state by year-quarter fixed effects are included and standard errors are double clustered at the state and year-quarter level. T-statistics are reported in the parentheses. The symbols *, **, and *** represent statistical significance at the 10%, 5%, and 1% levels, respectively.

| | CRA | | FDICIA | |
|---|---|---|---|---|
| | (1) | (2) | (3) | (4) |
| Time Trend x CRA | 0.018 | 0.024 | | |
| | (0.55) | (0.62) | | |
| Time Trend x FDICIA | | | 0.010 | 0.029 |
| | | | (0.22) | (0.49) |
| Constant | 0.253*** | 0.250*** | 0.253*** | 0.313*** |
| | (436.24) | (543.40) | (876.78) | (94.03) |
| Bank FE | Yes | Yes | Yes | Yes |
| State x Year-Quarter FE | Yes | Yes | Yes | Yes |
| Reg. Overlap | Yes | No | Yes | No |
| Observations | 117509 | 112811 | 117509 | 13028 |
| Adj. R2 | 0.924 | 0.921 | 0.924 | 0.957 |

Table 2.5: Liquidity Creation

This table shows the results of the regulatory changes to the CRA and FDICIA on liquidity creation. *Post* is a dummy variable equal to 1 for years after 2005 and zero for years prior to 2005. *CRA* is an indicator variable equal to 1 if the bank had assets above $250 million at the beginning of the current year and previous year, and *FDICIA* is an indicator variable equal to 1 if the bank had assets above $500 million at the beginning of the current year. Columns 1, 2, 5, and 6 use the full sample. Columns 3 and 4 exclude banks required to comply with FDICIA. Columns 7 and 8 exclude banks not required to comply with CRA. Size is the natural log of total assets. Risk based capital ratio is equity/total risk weighted assets. ROA is the quarterly net income/total assets. Non-performing loans is the sum of non-accrual loans and loans past 90 days overdue scaled by total loans. The control variables are lagged. Bank and state by year-quarter fixed effects are included and standard errors are double clustered at the state and year-quarter level. The symbols *, **, and *** represent statistical significance at the 10%, 5%, and 1% levels, respectively.

| | CRA | | | | FDICIA | | | |
|---|---|---|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) |
| Post x CRA | 0.013*** | 0.016*** | 0.011** | 0.014*** | | | | |
| | (3.11) | (4.20) | (2.36) | (3.33) | | | | |
| Post x FDICIA | | | | | 0.018*** | 0.019*** | 0.009 | 0.006 |
| | | | | | (3.24) | (4.05) | (1.59) | (1.29) |
| Size | | -0.022*** | | -0.021*** | | -0.023*** | | -0.083*** |
| | | (-3.19) | | (-3.01) | | (-3.27) | | (-9.19) |
| Risk Based Cap. Ratio | | -0.930*** | | -0.926*** | | -0.929*** | | -1.154*** |
| | | (-27.96) | | (-27.60) | | (-27.95) | | (-12.80) |
| ROA | | 1.112*** | | 1.077*** | | 1.111*** | | 1.612*** |
| | | (7.63) | | (7.15) | | (7.63) | | (6.32) |
| Non-Performing Loans | | -0.301*** | | -0.291*** | | -0.298*** | | -0.429*** |
| | | (-6.57) | | (-6.27) | | (-6.59) | | (-4.50) |
| Constant | 0.276*** | 0.531*** | 0.271*** | 0.519*** | 0.276*** | 0.534*** | 0.343*** | 1.012*** |
| | (772.95) | (14.54) | (1098.59) | (14.54) | (1622.05) | (14.74) | (284.17) | (16.22) |
| Bank FE | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| State x Year-Quarter FE | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Reg. Overlap | Yes | Yes | No | No | Yes | Yes | No | No |
| Observations | 224506 | 224506 | 213047 | 213047 | 224506 | 224506 | 31505 | 31505 |
| Adj. R2 | 0.887 | 0.906 | 0.883 | 0.903 | 0.887 | 0.906 | 0.922 | 0.937 |

Table 2.6: Liquidity Creation Components

This table shows the results of the regulatory change to CRA reporting on the components of liquidity creation. *Post* is a dummy variable equal to 1 for years after 2005 and zero for years prior to 2005. *CRA* is an indicator variable equal to 1 if the bank had assets above $250 million at the beginning of the current year and previous year. Size is the natural log of total assets. Risk based capital ratio is equity/total risk weighted assets. ROA is the quarterly net income/total assets. Non-performing loans is the sum of non-accrual loans and loans past 90 days overdue scaled by total loans. The control variables are lagged. Bank and state by year-quarter fixed effects are included and standard errors are double clustered at the state and year-quarter level. The symbols *, **, and *** represent statistical significance at the 10%, 5%, and 1% levels, respectively.

| | (1) Illiquid Assets | (2) Illiquid Liabilities | (3) Liquid Assets | (4) Liquid Liabilities | (5) Illiquid Guarantees |
|---|---|---|---|---|---|
| Post x CRA | 0.011*** | 0.001 | -0.006* | 0.008** | 0.004* |
| | (3.28) | (1.06) | (-1.85) | (2.43) | (1.73) |
| Size | 0.016** | -0.006*** | 0.007 | -0.054*** | -0.006** |
| | (2.29) | (-6.16) | (1.12) | (-13.58) | (-2.03) |
| Risk Based Cap. Ratio | -0.602*** | 0.297*** | 0.834*** | -0.020 | -0.093*** |
| | (-18.01) | (26.47) | (24.06) | (-0.81) | (-7.41) |
| ROA | 0.611*** | 0.217*** | -0.880*** | 0.361*** | 0.600*** |
| | (4.61) | (6.58) | (-5.91) | (4.65) | (7.59) |
| Non-Performing Loans | 0.013 | -0.032*** | 0.084* | -0.190*** | -0.333*** |
| | (0.28) | (-3.71) | (1.79) | (-5.72) | (-13.10) |
| Constant | 0.468*** | 0.088*** | 0.158*** | 0.688*** | 0.138*** |
| | (13.30) | (16.89) | (5.13) | (34.79) | (9.48) |
| Bank FE | Yes | Yes | Yes | Yes | Yes |
| State x Year-Quarter FE | Yes | Yes | Yes | Yes | Yes |
| Observations | 213047 | 213047 | 213047 | 213047 | 213047 |
| Adj. R2 | 0.909 | 0.861 | 0.871 | 0.862 | 0.788 |

Table 2.7: Small Business and Farm Lending

This table shows the results of the regulatory change to CRA reporting on small business lending. Specifically this looks at the totals of non-farm and farm related loans as well as the overall small business lending. Column 1 looks at the number of small business (non farm) loans outstanding, column 2 looks at the dollar volume outstanding of small business (non farm) loans outstanding, column 3 looks at the number of agricultural loans outstanding, column 4 looks at the dollar volume outstanding of agricultural outstanding, column 5 looks at the overall number of small business loans outstanding, and column 6 looks at the overall dollar volume outstanding of small business loans outstanding. *Post* is a dummy variable equal to 1 for years after 2005 and zero for years prior to 2005. *CRA* is an indicator variable equal to 1 if the bank had assets above $250 million at the beginning of the current year and previous year. Size is the natural log of total assets. Risk based capital ratio is equity/total risk weighted assets. ROA is the quarterly net income/total assets. Non-performing loans is the sum of non-accrual loans and loans past 90 days overdue scaled by total loans. The control variables are lagged. Bank and state by year fixed effects are included and standard errors are double clustered at the state and year level. The symbols *, **, and *** represent statistical significance at the 10%, 5%, and 1% levels, respectively.

| | Small Business | | Small Farm | | Total | |
|---|---|---|---|---|---|---|
| | (1) No. Loans | (2) $ Outstanding | (3) No. Loans | (4) $ Outstanding | (5) No. Loans | (6) $ Outstanding |
| Post x CRA | -18.123 | 4.118*** | -5.847 | 0.332 | -23.970 | 4.449*** |
| | (-1.35) | (4.74) | (-1.13) | (1.37) | (-1.44) | (4.37) |
| Size | 153.607*** | 18.233*** | 55.623*** | 3.403*** | 209.230*** | 21.635*** |
| | (10.98) | (18.55) | (5.35) | (4.77) | (11.79) | (20.77) |
| Risk Based Cap. Ratio | -247.745*** | -23.236*** | -36.805* | -3.626*** | -284.550*** | -26.862*** |
| | (-8.39) | (-9.16) | (-1.87) | (-3.71) | (-6.78) | (-10.47) |
| ROA | 1022.365*** | 51.106*** | 313.769** | 18.507*** | 1336.134*** | 69.613*** |
| | (4.69) | (3.59) | (2.96) | (3.53) | (4.52) | (4.29) |
| Non-Performing Loans | -179.890** | -25.081*** | -37.001 | -9.682*** | -216.891** | -34.763*** |
| | (-3.08) | (-5.01) | (-1.19) | (-5.41) | (-2.84) | (-6.70) |
| Constant | -372.161*** | -55.689*** | -96.163* | -7.477** | -468.324*** | -63.167*** |
| | (-5.84) | (-12.18) | (-1.98) | (-2.32) | (-5.70) | (-12.83) |
| Bank FE | Yes | Yes | Yes | Yes | Yes | Yes |
| State x Year FE | Yes | Yes | Yes | Yes | Yes | Yes |
| Observations | 52991 | 52991 | 52991 | 52991 | 52991 | 52991 |
| Adj. R2 | 0.838 | 0.912 | 0.911 | 0.938 | 0.864 | 0.912 |

## Table 2.8: SBA Lending

This table shows the results of the regulatory change to CRA reporting related to SBA guaranteed loans in CRA-eligible and non-CRA-eligible zip codes. The group is *Post* is a dummy variable equal to 1 for years after 2005 and zero for years prior to 2005. The year 2005 is excluded as the regulation change was announced and implemented during that year. *CRA* is an indicator variable equal to 1 if the bank had assets above $250 million at the beginning of the current year and previous year. Size is the natural log of total assets. Risk based capital ratio is equity/total risk weighted assets. ROA is the quarterly net income/total assets. Non-performing loans is the sum of non-accrual loans and loans past 90 days overdue scaled by total loans. The control variables are lagged. Bank and state by year fixed effects are included and standard errors are clustered at the year level. The symbols *, **, and *** represent statistical significance at the 10%, 5%, and 1% levels, respectively.

| | CRA-Eligible | | | | Not CRA-Eligible | | | |
|---|---|---|---|---|---|---|---|---|
| | (1) Number Loans | (2) Loan Size | (3) Jobs Supported | (4) % Loan Failure | (5) Number Loans | (6) Loan Size | (7) Jobs Supported | (8) % Loan Failure |
| Post x CRA | -0.621 | -0.090 | -2.577 | 0.027 | 0.259 | 0.213 | 27.796*** | -0.025 |
| | (-1.51) | (-0.35) | (-0.49) | (0.33) | (0.65) | (1.24) | (3.33) | (-1.28) |
| Size | 1.057 | 0.505 | 39.819** | -0.054 | 1.461** | 0.602*** | 40.603*** | -0.011 |
| | (1.53) | (1.53) | (2.97) | (-0.61) | (3.14) | (3.95) | (4.76) | (-0.36) |
| Risk Based Cap. Ratio | 7.562** | 5.046** | 178.544*** | 0.038 | 1.978 | 2.171* | 82.448** | 0.013 |
| | (2.24) | (2.94) | (5.17) | (0.11) | (1.08) | (1.95) | (2.51) | (0.08) |
| ROA | 9.652 | -7.149 | -96.093 | 2.681 | 30.814* | 7.481* | 318.969 | 0.769 |
| | (0.45) | (-1.29) | (-0.42) | (1.48) | (2.19) | (1.90) | (1.42) | (1.07) |
| Non-Performing Loans | 11.019* | -1.736 | 89.128 | -1.393 | -0.801 | -1.517 | -3.510 | -0.445 |
| | (1.82) | (-0.26) | (0.45) | (-1.15) | (-0.12) | (-0.55) | (-0.02) | (-1.25) |
| Constant | -4.352 | -2.277 | -197.651** | 0.402 | -4.146 | -2.181** | -180.870*** | 0.213 |
| | (-1.21) | (-1.42) | (-2.86) | (0.85) | (-1.70) | (-2.69) | (-4.04) | (1.39) |
| Bank FE | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| State x Year FE | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Observations | 1520 | 1520 | 1520 | 1378 | 8456 | 8456 | 8456 | 7923 |
| Adj. R2 | 0.656 | 0.583 | 0.387 | 0.019 | 0.786 | 0.660 | 0.597 | 0.105 |

Table 2.9: Credit Risk: Past Due Loans

This table shows the results of the regulatory change to CRA reporting on past due loans. *Post* is a dummy variable equal to 1 for years after 2005 and zero for years prior to 2005. *CRA* is an indicator variable equal to 1 if the bank had assets above $250 million at the beginning of the current year and previous year. Size is the natural log of total assets. Risk based capital ratio is equity/total risk weighted assets. ROA is the quarterly net income/total assets. The control variables are lagged. Bank and state by year-quarter fixed effects are included and standard errors are double clustered at the state and year-quarter level. The symbols *, **, and *** represent statistical significance at the 10%, 5%, and 1% levels, respectively.

| | (1) | (2) | (3) | (4) |
|---|---|---|---|---|
| | Non-Performing Loans | Past Due Loans 30-89 | Past Due Loans 90+ | Non-Accrual Loans |
| Post x CRA | 0.002** | 0.000 | 0.000 | 0.002** |
| | (2.59) | (1.15) | (1.62) | (2.57) |
| Size | 0.006*** | 0.003*** | 0.001*** | 0.004*** |
| | (6.70) | (5.92) | (5.40) | (6.47) |
| Risk Based Cap. Ratio | 0.006** | 0.001 | -0.001 | 0.007*** |
| | (2.39) | (0.40) | (-0.89) | (3.05) |
| ROA | -0.787*** | -0.167*** | -0.102*** | -0.680*** |
| | (-11.73) | (-7.74) | (-6.92) | (-12.29) |
| Constant | -0.006 | 0.001 | -0.001 | -0.005* |
| | (-1.62) | (0.52) | (-0.91) | (-1.77) |
| Bank FE | Yes | Yes | Yes | Yes |
| State x Year-Quarter FE | Yes | Yes | Yes | Yes |
| Observations | 213035 | 196947 | 213035 | 213035 |
| Adj. R2 | 0.533 | 0.475 | 0.357 | 0.512 |

Table 2.10: Credit Risk: Charge Offs

This table shows the results of the regulatory change to CRA reporting on loan charge offs and recoveries. *Post* is a dummy variable equal to 1 for years after 2005 and zero for years prior to 2005. *CRA* is an indicator variable equal to 1 if the bank had assets above $250 million at the beginning of the current year and previous year. Size is the natural log of total assets. Risk based capital ratio is equity/total risk weighted assets. ROA is the quarterly net income/total assets. The control variables are lagged. Bank and state by year-quarter fixed effects are included and standard errors are double clustered at the state and year-quarter level. The symbols *, **, and *** represent statistical significance at the 10%, 5%, and 1% levels, respectively.

| | (1) Charge-Offs | (2) Recoveries | (3) Net Charge-Offs |
|---|---|---|---|
| Post x CRA | 0.000 | -0.000 | 0.000 |
| | (1.03) | (-0.14) | (1.12) |
| Size | 0.002*** | -0.000 | 0.002*** |
| | (6.43) | (-0.17) | (6.96) |
| Risk Based Cap. Ratio | 0.003** | 0.003*** | 0.001 |
| | (2.43) | (9.86) | (0.57) |
| ROA | -0.544*** | -0.026*** | -0.512*** |
| | (-14.98) | (-12.61) | (-14.99) |
| Constant | -0.001 | 0.001*** | -0.002 |
| | (-0.88) | (2.99) | (-1.44) |
| Bank FE | Yes | Yes | Yes |
| State x Year-Quarter FE | Yes | Yes | Yes |
| Observations | 213035 | 213035 | 213035 |
| Adj. R2 | 0.554 | 0.453 | 0.530 |

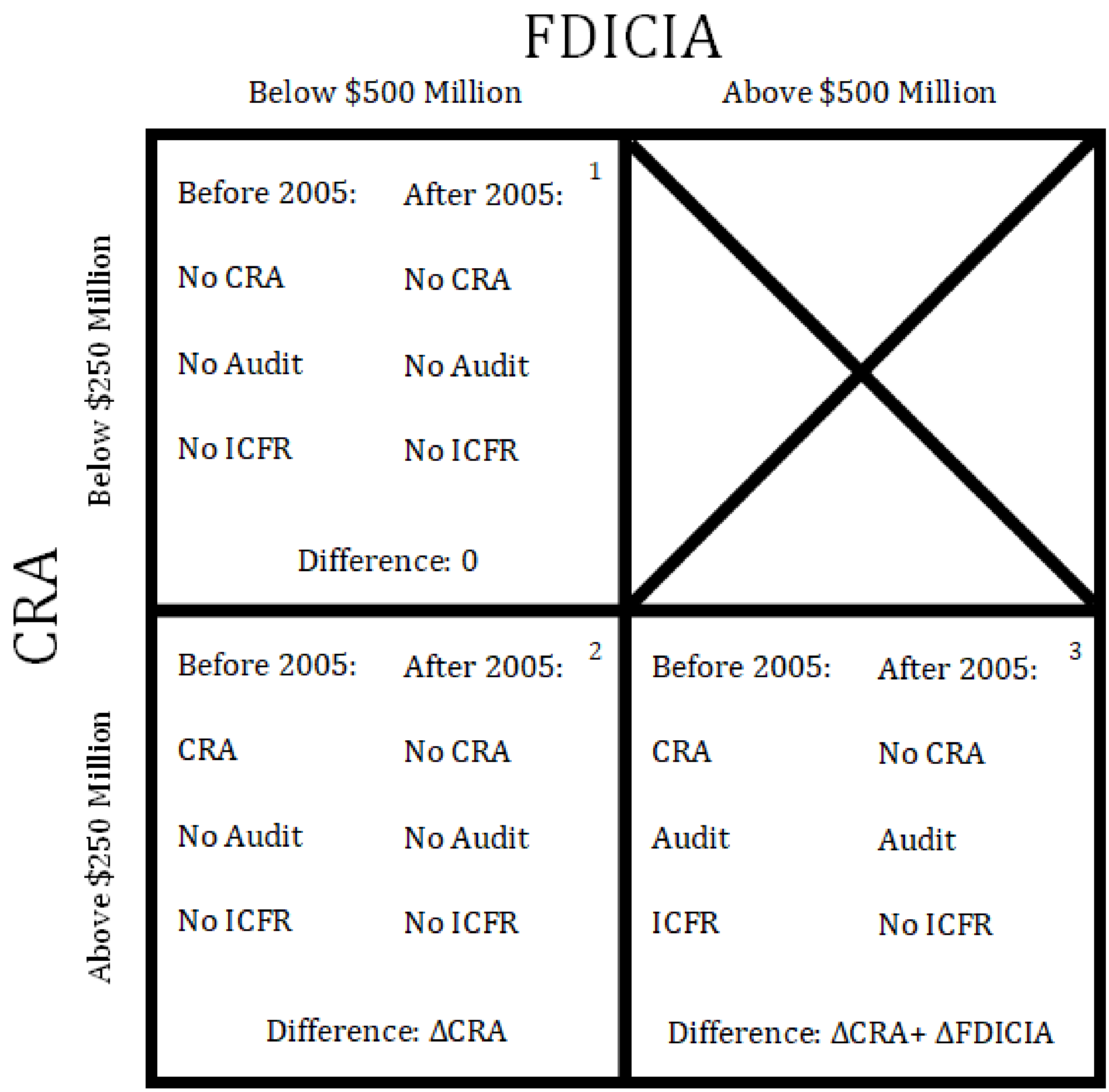



Figure 2.1: Regulatory Threshold Overlap

This figure shows the size-based threshold for the CRA and FDICIA as well as provides a graphical representation of the overlap between the two regulations. The figure shows the size-based threshold for both the CRA and FDICIA regulations. It also shows the regulatory requirements associated with the two regulations before and after the 2005 regulatory change. The first difference of each group is also provided.

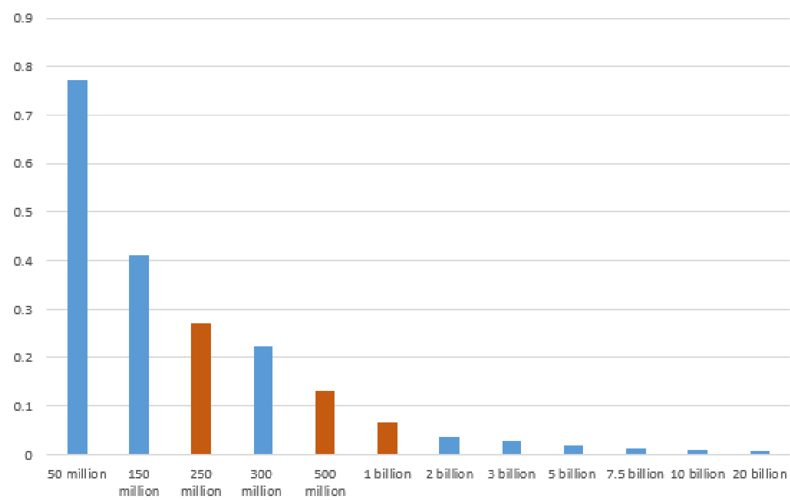



Figure 2.2: Percent of Banks at various thresholds

This figure shows the percentage of the number of financial institutions as each threshold provided in Table 2.1 as of December 31, 2005. This only includes thresholds that have bank specific regulatory size thresholds. In addition, some thresholds did not exist in 2005 (such as the one related to TILA), so the earliest threshold level is included.

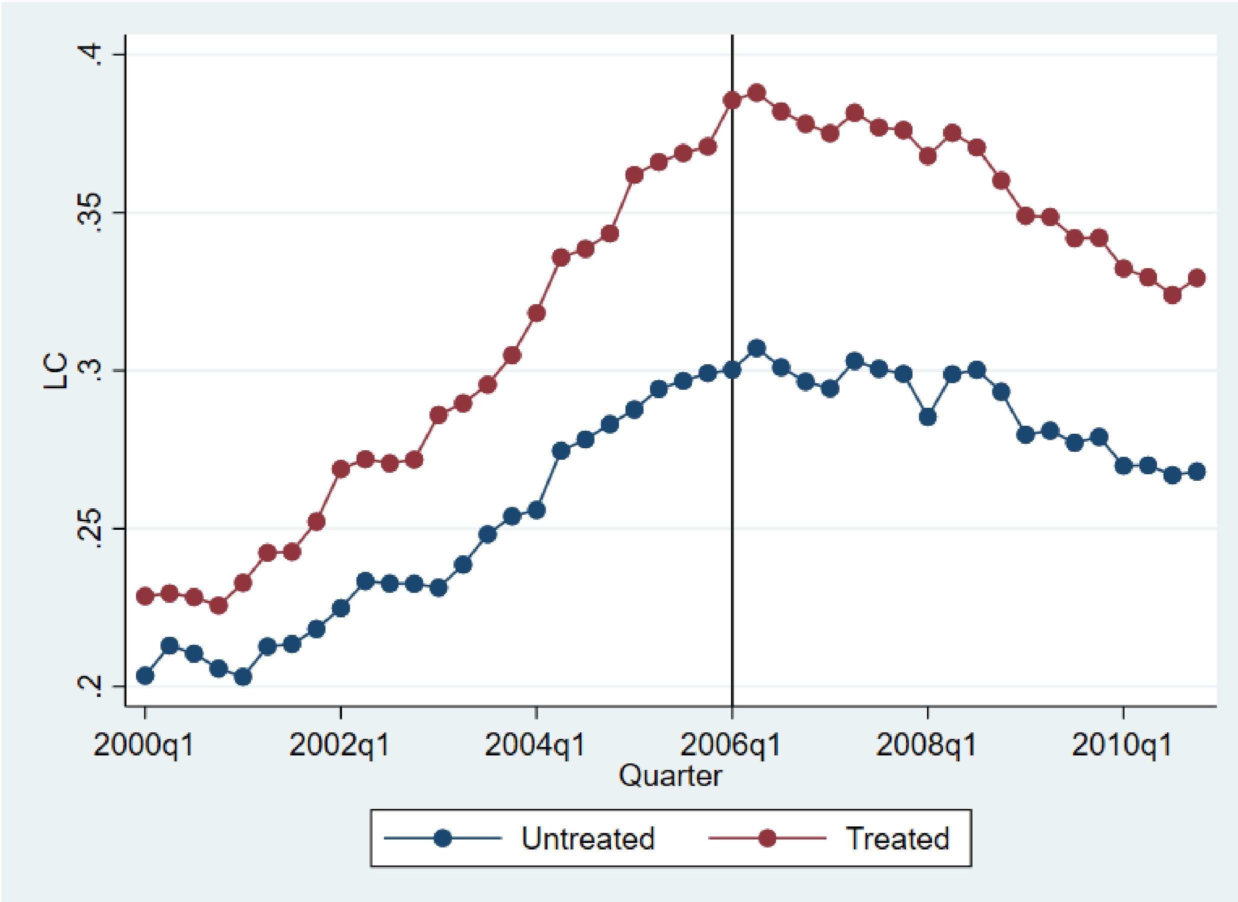



Figure 2.3: Univariate Trend Analysis: CRA

This figure plots the average value of LC(Total) by quarter for the period of analysis for the treated and control banks for the CRA regulatory change. The vertical line represents the delineation between the pre-regulatory and post-regulatory change.

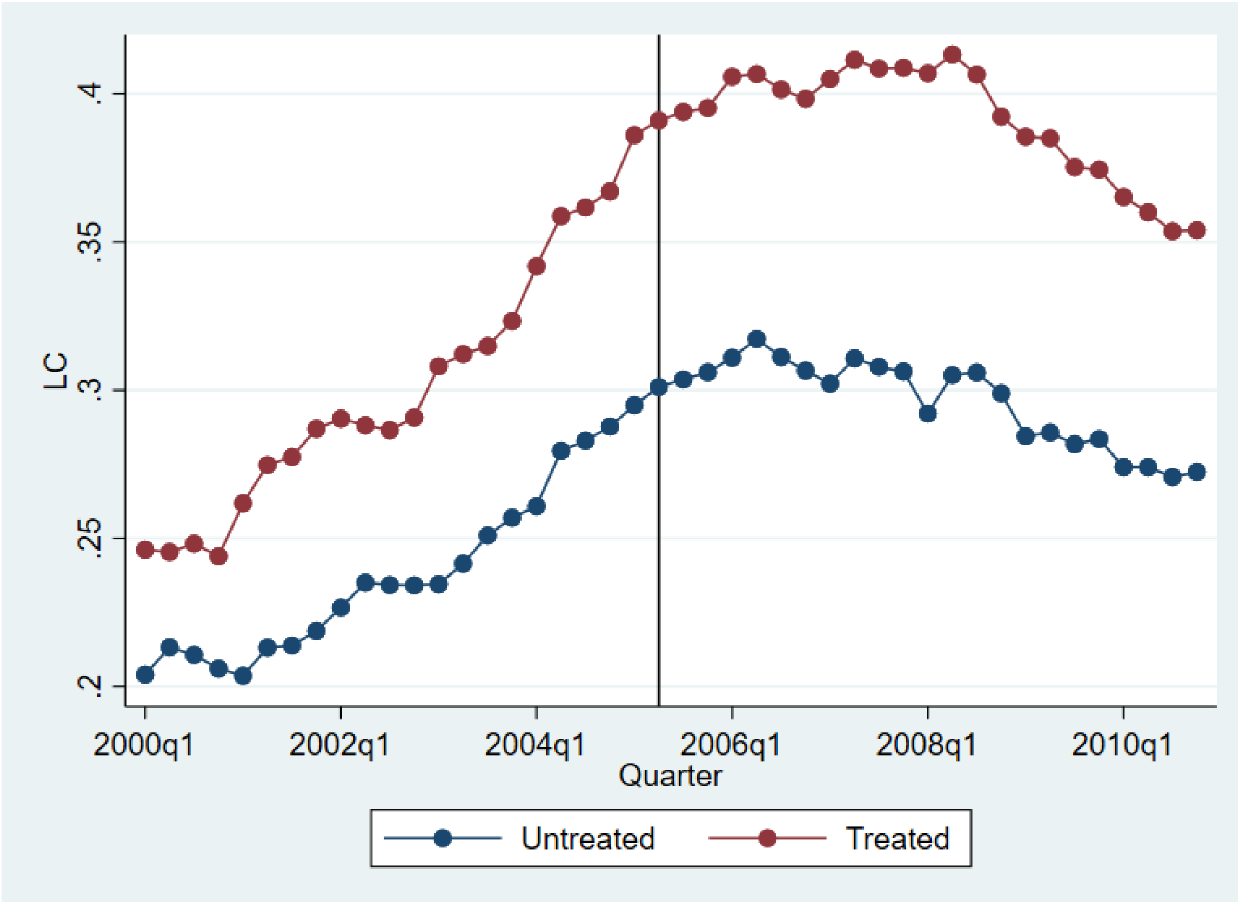
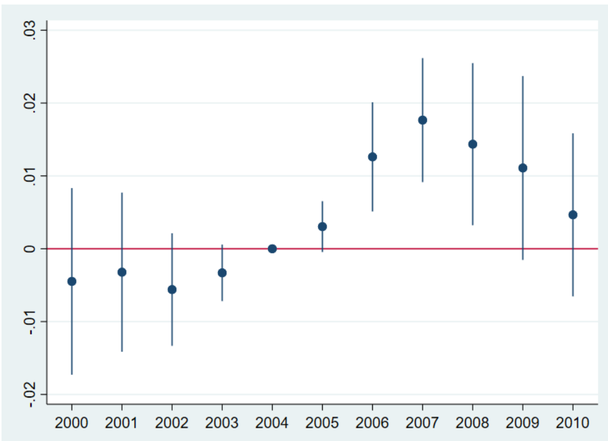
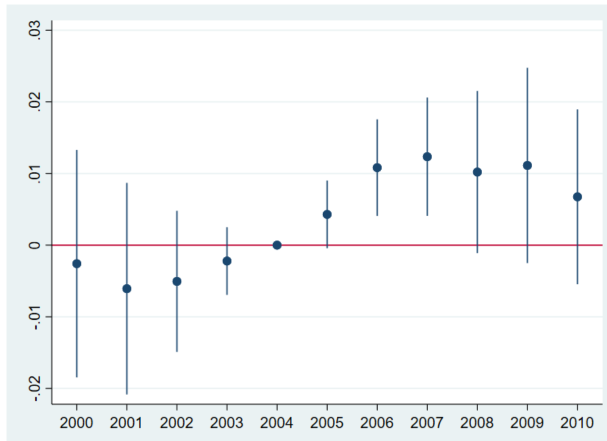



Figure 2.4: Univariate Trend Analysis: FDICIA

This figure plots the average value of LC(Total) by quarter for the period of analysis for the treated and control banks for the FDICIA regulatory change. The vertical line represents the delineation between the pre-regulatory and post-regulatory change.



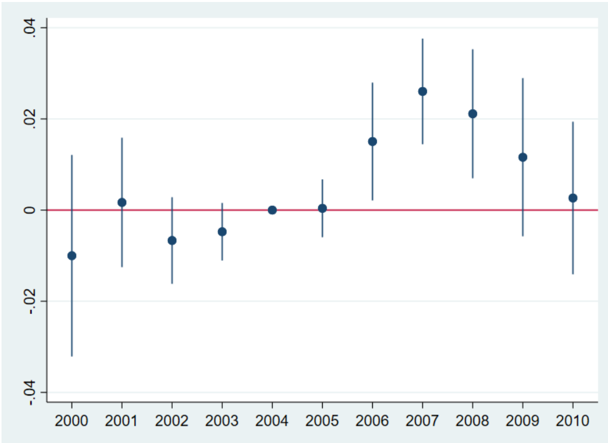

(a) Full Sample



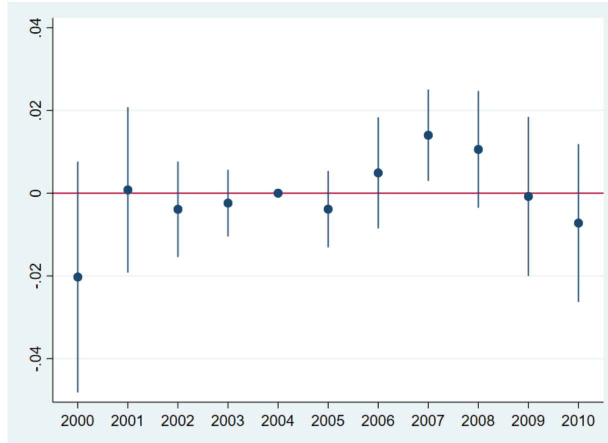

(b) Preferred Sample

Figure 2.5: Parallel Trend Analysis: CRA

This figure plots coefficients of the interaction term in equation 2.7 for the FDICIA regulatory change. Observations are quarterly observations by bank though the analysis is run using an interaction between the treatment indicator $CRA$ and $\delta_t$ which an indicator for a given year. The year 2005 is excluded from the analysis and 2004 is treated as the excluded year. Bank and state by year fixed effects are included and standard errors are double clustered at the bank and year level. In the left figure, I plot the coefficients for the analysis where each regulation is treated separately. In the right figure, I plot the coefficients for my preferred sub-sample that addresses the regulatory overlap.




(a) Full Sample




(b) Preferred Sample

Figure 2.6: Parallel Trend Analysis: FDICIA

This figure plots coefficients of the interaction term in equation 2.7 for the FDICIA regulatory change. Observations are quarterly observations by bank though the analysis is run using an interaction between the treatment indicator *FDICIA* and $\delta_t$ which an indicator for a given year. The year 2005 is excluded from the analysis and 2004 is treated as the excluded year. Bank and state by year fixed effects are included and standard errors are double clustered at the bank and year level. In the left figure, I plot the coefficients for the analysis where each regulation is treated separately. In the right figure, I plot the coefficients for my preferred sub-sample that addresses the regulatory overlap.

# Chapter 3

# Uncertainty and The Rationality of Market Earnings Expectations [1]

We thoroughly examine the relationship between uncertainty and the rationality of market earnings expectations, using multiple uncertainty measures and the return predictability of analysts' conditional biases (EHB). Contrary to predictions of existing theories, we find a positive time-series relationship and a negative cross-sectional relationship. To explain these results, we propose a limited attention model where information processing costs increase with uncertainty. Our model offers distinct predictions for investors' attention allocation and information acquisition compared to existing models. We find strong empirical support for

[1] Co-Authors: Harrison Ham (University of Georgia), Zhongjin (Gene) Lu (University of Georgia), Wang Renxuan (China Europe International Business School)

our model's predictions using direct measures of investor attention. Our model also reconciles seemingly contradictory existing findings.

## 3.1 Introduction

Is the stock market more or less rational when uncertainty is high? Despite decades long research into financial markets, we have surprisingly little evidence to understand this important question. In fact, existing theories provide conflicting perspectives on this issue. On one hand, behavioral models posit that market expectations may become less rational when uncertainty is high, due to a positive correlation between uncertainty and investors' behavioral biases (e.g., Hirshleifer, 2001). Conversely, information choice models predict that high uncertainty increases the value of information, leading to increased information acquisition and more rational market expectations (e.g., see the survey in Baley and Veldkamp, 2021; Mackowiak et al., 2021).

In this paper, we provide insight into this question by studying the relationship between uncertainty and the degree of rationality in the market's earnings expectations. Our investigation builds on a literature that shows analysts' forecasts play an important role in shaping the market's earnings expectations (Kothari et al., 2016), and that the market is not fully rational in processing information in analysts' forecasts, as evidenced by the return predictability of ex-ante predictable biases in analysts' earnings forecasts (Frankel and Lee, 1998; So, 2013; J. van Binsbergen et al., 2020). Furthermore, we take advantage of the recent advancement in measuring these biases, defined as the real-time difference between analysts' forecasts and the optimal statistical forecasts generated by machine-learning algorithms (J. van Binsbergen et al., 2020). We refer to this empirical measure as ex-ante human bias (hereafter, EHB), and investigate how the degree of return predictability of EHB relates to different levels of uncertainty.[2]

Inconsistent with the predictions of either of the theories, we find a contrasting relationship between uncertainty and market rationality in the time-series versus the cross-section.[3] Specifically, we find the return predictability of analysts' conditional bias being *stronger* among firms with higher levels of uncertainty, but *weaker* during periods of higher uncertainty. This contrasting relationship is visualized in Figure 3.1. As shown in this figure, this opposite pattern holds regardless of the use of realized volatility or option-implied volatility. It is also not driven by the difference between financial volatility and fundamental volatility nor the difference between firm-level volatility and market-level volatility . This contrasting relationship thus presents a challenge for existing theories, as standard behavioral stories can only explain the positive cross-sectional relation but not the negative time-series relation, whereas the opposite is true for standard information choice models.

We propose to explain our finding by hypothesizing that investors' information processing costs to unravel analysts' biases increase with the level of uncertainty. Here, higher information processing costs mean investors need to deploy more effort/capacity to process the same

[2]Dávila and Parlatore, 2023 construct a measure for the comovement between price informativeness and volatility, while our measure focuses directly on the rationality of market earnings expectations with respect to the efficiency in processing analysts' forecasts.

[3]All our empirical results focus on non-microcap stocks, defined as stocks larger than the 20th percentile of the NYSE size distribution to avoid the disproportional influence from small and illiquid stocks in a regression.

amount of information in order to unravel analysts' biases.[4] The hypothesis is intuitive for both the cross-sectional and time-series variation of uncertainty. It simply takes more effort and mental capacity to understand the earnings of a young pharmaceutical company, which has less information available and more uncertain future prospects than a mature industrial firm. Similarly, analyzing the same company becomes more demanding during uncertain times, such as the Great Financial Crisis than during the peaceful periods preceded 2007. We term this hypothesis the "information cost channel".

We incorporate the information cost channel into a standard information choice model. We show that the strength of the information cost channel – measured by the correlation between information costs and uncertainty – plays a critical role in driving the relation between uncertainty and the rationality of market earnings expectations. When the information cost channel is weak, as in the existing canonical information choice models, the relation between uncertainty and the rationality of market earnings expectations is unambiguously positive. When the information cost channel is strong, this relation can turn negative.

Our models leads to four sets of interesting predictions. First, our model offers a rational explanation of the long-lasting return predictability of EHB. Since debiasing analysts' forecasts is costly, investors unravel the bias fully only when the financial stake is sufficiently high. Consistent with this prediction, we find that the return predictability of EHB is statis-

[4]Sims, 2003 is the seminal limited attention paper. See, for example, Blankespoor et al., 2020 for a review of the information processing cost literature and L. Veldkamp, 2011 for a the textbook treatment of the limited-inattention theories.

tically significant only among small- or medium-sized firms and insignificant among mega-cap stocks.[5]

Second, since information processing costs is determined by the information environment, which varies substantially across firms but is unlikely to vary at a high frequency, our model predicts heterogeneous and potentially even opposite relationships in the cross-section versus the time-series. Our finding of opposite relationships in the cross-section versus time-series therefore confirms the important role the information cost channel may play in shaping market rationality, which have been largely ignored in the literature.

In our third set of tests, we use a firm's age as a measure of information costs to directly test the presence of the information cost channel. Intuitively, younger firms simply have less information available to investors and therefore should require greater costs to obtain an additional unit of information. Our model predicts that the return predictability of EHB will weaken as firms age. Furthermore, the model predicts that as a firm's information environment improves, thus lowering information processing costs, the return predictability of EHB should become weaker. Through sorting firms into portfolios based on their age, we find stronger return predictability among the younger firms. Furthermore, we find the return predictability to significantly decrease when EHB after firms' earnings announcements, when the information environment improves.

Fourth, we derive the model's prediction on investors' attention allocation with and without the information cost channel. To that end, we find canonical models, where information

[5]The small-, medium-, and mega-cap stocks are defined as stocks between the 20th and 50th percentiles, 50th and 80th percentiles, and above 80th percentile of the NYSE size distribution. The mega-cap stocks account for almost 80% of the total market capitalization.

costs do not vary with uncertainty, lead to very different predictions about how investors allocate their attention when facing uncertainty spikes compared to our model with information costs. Specifically, absent of information cost channel, investors allocate more attention for firms with lower uncertainty. This is because investors have already placed greater attention on highly uncertain firms. Therefore the additional benefit from further attention allocation is less for these firms. On the other hand, the opposite holds true when the information cost channel is strong — investors have allocated relatively less attention to highly uncertain firms due to the higher information costs. Therefore, as time-series uncertainty spikes, the benefit of paying attention to more uncertain firms is greater than those with less uncertainty. Using the abnormal institutional attention from Bloomberg Ben-Rephael et al., 2017, hereafter AIA), we find strong evidence supporting the presence of a strong information cost channel.

Finally, our model helps reconcile existing findings regarding how market expectations incorporate analysts' forecasts when facing uncertainty. Zhang, 2006 find investors react *less* completely to analysts' forecasts among firms with high uncertainty, whereas Loh and Stulz, 2018 investors rely *more* on analysts' forecasts when the aggregate uncertainty or the firm-level uncertainty is high. Since the information cost channel is strong for cross-sectional variation in uncertainty, our model provides an alternative explanation to Zhang, 2006's finding based on limited attention rather than behavioral bias. In our model, investors endogenously choose to place a higher weight on analysts' forecasts when facing increased uncertainty (regardless of aggregate or firm-level), leading to Loh and Stulz, 2018 's findings; at the same time, investors choose to debias less among firms with high uncertainty due to the information cost channel, leading to Zhang, 2006's finding.

Overall, our theoretical and empirical results contribute to the limited attention literature by highlighting the effect of uncertainty on the information cost, which has been largely ignored.[6] To the best of our knowledge, our study is also the first to use the limited attention theory to explain the relation between uncertainty and the rationality of market earnings expectations.[7]

Our findings complement two recent studies on the impact of uncertainty on investors' attention and price efficiency.[8] Andrei et al., 2021 study how the market-level uncertainty affects investors' attention to earnings announcements, finding that VIX is positively related to investors' attention and the strength of price reaction to earnings announcements. Benamar et al., 2021 use novel click data to measure uncertainty and find a positive time-series relation between uncertainty and the reaction of U.S. Treasury price to macroeconomic announcements. The findings of these two studies support the information benefit channel, which has been emphasized in theoretical studies (e.g., see the survey in Baley and Veldkamp, 2021; Mackowiak et al., 2021). We contribute to the literature by highlighting that the information benefit channel alone is not sufficient to fully explain the relationship between uncertainty and price efficiency as measured by the return predictability of EHB. We show the information cost channel, which allows for the positive correlation between uncertainty

[6] In Section 2, we discuss why the effect of uncertainty on the information benefit dominates in effect of uncertainty on the information cost in canonical information choice models.

[7] The limited attention has been used in finance literature to explain the home bias puzzle (Van Nieuwerburgh and Veldkamp, 2009), the under diversification puzzle (Van Nieuwerburgh and Veldkamp, 2010), asset return comovements (Peng and Xiong, 2006; L. L. Veldkamp, 2006). The theory has also been applied in other fields of economics, as surveyed in Mackowiak et al., 2021.

[8] The empirical studies on investors' attention have established a robust, positive link between investors' attention and stock price efficiency (e.g., Ben-Rephael et al., 2017; Da et al., 2011; Dellavigna and Pollet, 2009; Hirshleifer et al., 2009; S. Huang et al., 2019).

and information processing cost, is needed for the limited attention theory to fully explain the variations in the return predictability of EHB.[9]

Finally, our study makes a contribution to the literature on stock prices and analysts' forecasts (Frankel and Lee, 1998; So, 2013; J. van Binsbergen et al., 2020) by proposing an information choice model to explain the puzzle of why investors do not fully account for predictable analysts' forecast errors, which has been a longstanding issue in the literature (see the survey in Kothari et al., 2016). Our model also reveals novel determinants of the return predictability of EHB, which we show empirically explain large variations in the return predictability of EHB.

## 3.2 Theoretical Framework

We tailor a standard information choice model environment to our setting, in which analysts' forecasts are an important information source for the market's earnings expectations and investors need to debias analysts forecasts to obtain more precise signals for future earnings. We show that investors' optimal attention allocation and the extent to which they debias analysts forecasts are endogenously linked to the marginal benefit and cost of information. Most importantly, we define the information cost channel and derive model predictions with and without the information cost channel to inform our empirical tests.

[9] Dávila and Parlatore, 2023 show that volatility and price informativeness can comove positively or negatively depending on the primitives of the model environment. They define price informativeness as the precision of the unbiased signal of the innovation to the asset payoff. While related, price informativeness is distinct from price efficiency. For example, in a world in which price reflects all available information but is heavily influenced by noise (e.g., unpredictable shocks to sentiment), price informativeness can be low while price is fully efficient.

### 3.2.1 Investors' Information Environment

The rational inattention framework (Sims, 2003) has been widely used in the macroeconomics and finance literature (see Baley and Veldkamp, 2021; Mackowiak et al., 2021 for excellent reviews). We apply this theoretical framework to model how investors unravel the conditional biases in analysts' forecasts.

A firms' earnings $y$ is drawn independently from a normal distribution $N\left(0, \tau_0^{-1}\right)$, where$\tau_0^{-1}$ captures the degrees of fundamental uncertainty investors face. Analysts' forecasts ($AF$) are the sum of the private signal ($\theta$) and the bias ($B$):

$$AF = \theta + B,$$

where the signal $\theta = y + \eta$ with $\eta$ denoting the noise in the signal, which is independently drawn from $N\left(0, \tau_\eta^{-1}\right)$.[10] Under the assumption that $B$ is uncorrelated with $\theta$, we have $\tau_{AF}^{-1} = \tau_\eta^{-1} + \tau_B^{-1}$.

There is no cost for investors to use $AF$ at face value to update their earnings expectations, but if investors want to unravel the conditional biases in $AF$, they need to pay for the costly information-processing capacity.

The degree to which they unravel analysts' biases is an endogenous outcome of the trade-off between obtaining accurate information and the costs of processing that information. Specifically, we denote the degree to which investors unravel analysts' conditional bias as $b$

[10] Existing literature discusses many incentive reasons why $B$ is not zero and shows that $B$ is ex ante measurable. See Kothari et al., 2016 for a review.

and the resulting investors' signal after debiasing by $s$:

$$s = AF - b \times B$$

where $s \sim N\left(0, \tau_s^{-1}\right)$ with $\tau_s^{-1} = \tau_\eta^{-1} + \tau_B^{-1} \times (1-b)^2$. Therefore, when investors unravel more completely, $b$ increases toward 1, and the signal $s$ becomes more accurate. When $b = 1$ and $s = \theta$, investors fully unravel the bias and obtain the best forecasts.

### 3.2.2 Investors' optimization problem

We first derive the optimal allocation of investors' costly attention and the resulting signal precision for a given level of fundamental uncertainty $\tau_0^{-1}$. We then discuss the effects of changing uncertainty on these optimal solutions without the information cost channel, through the information benefit channel show how changes in uncertainty impact these optimal solutions.

Investors consider paying attention to debias analysts' forecasts costly, but more attention leads to more accurate signal, which reduces the resulting squared forecasting errors, a desirable outcome for investors. Following standard information choice models, we model investors' trade-off as follows:

$$\min_C \frac{r}{2}E[y - a(C)]^2 + C \tag{3.1}$$

$$s.t.\ \kappa = \frac{C}{\lambda}$$

$$0 \le \frac{1}{2}log\frac{\tau_s + \tau_0}{\tau_{AF} + \tau_0} \le \kappa, \tag{3.2}$$

where $r$ is investors' financial stake in this firm, $C$ is investors' attention, and $a(C)$ is investors' optimal forecasts given their attention to debias analysts' forecasts.

The information processing capacity ($\kappa$) is determined by investors' attention ($C$) and the unit cost of processing information ($\lambda$). That is, if a firm is more difficult to analyze (high $\lambda$), the same amount of attention will result in less information capacity ($\kappa$), or bring less information benefit. The information processing capacity constrains the maximum reduction in uncertainty (entropy) from paying attention to debias analysts' forecasts ($\frac{1}{2}log\frac{\tau_s+\tau_0}{\tau_{AF}+\tau_0}$). [11]

**Proposition 1.** *(Optimal Attention Allocation and Resulting Signal Precision)*

When $\frac{r}{\lambda} < \bar{c} = \tau_\eta + \tau_0$, the information constraint in Eq. (3.2) is binding, we have the following optimal solutions for the optimal attention ($C^*$) and information processing capacity ($\kappa^*$)

$$C^* = \frac{\lambda}{2}ln\left(\frac{r}{\lambda}\right) - \frac{\lambda}{2}ln\left(\frac{1}{\tau_{AF}^{-1}} + \frac{1}{\tau_0^{-1}}\right), \tag{3.4}$$

[11] Intuitively, investors will unravel the bias at their full information processing capacity, i.e., $\frac{\tau_s+\tau_0}{\tau_{AF}+\tau_0} = e^{2\kappa}$. Following Baley and Veldkamp, 2021, given $s$, the optimal earnings forecast for investors is

$$a^* = E\left(y_{i,t+1}|s\right) = \frac{\tau_s}{\tau_0 + \tau_s}s \tag{3.3}$$

The resulting expected loss is $E\left[(y-a)^2\right] = \frac{1}{\tau_0+\tau_s}$, which decreases as $\tau_s$ increases.

$$\kappa^* = \frac{1}{2}ln\left(\frac{r}{\lambda}\right) - \frac{1}{2}ln\left(\frac{1}{\tau_{AF}^{-1}} + \frac{1}{\tau_0^{-1}}\right) \tag{3.5}$$

Given the optimal attention allocation, the investor obtains a signal with the precision ($\tau_s^*$) equal to

$$\tau_s^* = e^{2\kappa^*}\left(\frac{1}{\tau_{AF}^{-1}} + \frac{1}{\tau_0^{-1}}\right) - \tau_0 = \frac{r}{\lambda} - \frac{1}{\tau_0^{-1}}, \tag{3.6}$$

and the extent of debiasing ($b^*$) analysts' forecasts is:

$$b^* = 1 - \sqrt{\frac{\tau_s^{*-1} - \tau_\eta^{-1}}{\tau_B^{-1}}} \tag{3.7}$$

When $\frac{r}{\lambda} > \bar{c} = \tau_\eta + \tau_0$, the information constraint in Eq. (3.2) is not binding. In this case, investors do not debias analysts' forecasts ($b^* = 1$).

*Proof.* See Appendix B.3. □

Equations (3.5) to (3.7) show that the optimal attention allocation and the resulting debiasing effort are functions of the financial stakes ($r$), the unit cost of processing information ($\lambda$) and uncertainty ($\tau_0^{-1}$ and $\tau_{AF}^{-1}$). We derive testable predictions based on the proposition about how investor attention and market rationality changes with respect to financial stakes, information costs, and uncertainty in Section 3.2.4

### 3.2.3 The Information Cost Channel

In the canonical information choice models, when facing higher uncertainty, limited attention investors need to weigh the benefits and costs of paying more attention. Intuitively, higher level of uncertainty increases the benefits of one additional unit of information, motivating investors to acquire more information. On the other hand, obtaining an additional unit of information is more costly when little is understood (the "familiarity" effect).[12] Overall, under the entropy-based information constraint, canonical information choice models predict the benefits over-weighs the costs, and investors should pay more attention to acquire more precise signals when forming expectations. We provide a more technical discussion on this point in Appendix B.3.

The key assumption these canonical models make is that the level of information cost ($\lambda$) itself does not vary with uncertainty. We argue this assumption should be relaxed: $\lambda$ could vary positively with uncertainty, which we refer to as the information cost channel, formally defined below.

**Definition.** Define the information cost channel as

$$\frac{\partial \lambda}{\partial \tau_0^{-1}} > 0 \tag{3.8}$$

i.e. the cost of information processing cost, $\lambda$ increases with the uncertainty of the underlying earnings process, $\tau_0^{-1}$.

[12]See p.21 of L. Veldkamp, 2011. We also thank Liyan Yang for sharing his discussion slides, which refer to this effect as the "familiarity" effect.

The information cost channel formalizes the intuition that processing unit information is more costly for uncertain states of the world: debiasing analysts forecasts for a young pharmaceutical company or during the global financial crisis simply takes more effort. With the information choice channel, the prediction of the information choice model regarding the relationship between uncertainty and market rationality changes.

**Proposition 2.** *Absent of the information cost channel ($\frac{\partial \lambda}{\partial \tau_0^{-1}} = 0$), the relation between uncertainty and the rationality of earnings expectations is* ***positive****. In the presence of the information choice channel ($\frac{\partial \lambda}{\partial \tau_0^{-1}} > 0$), this relation can turn* ***negative*** *when $\frac{\partial \lambda}{\partial \tau_0^{-1}}$ is sufficiently positive.*

*Proof.* See Appendix B.3. □

Proposition 2 states that in the presence of the information cost channel, i.e. when $\frac{\partial \lambda}{\partial \tau_0^{-1}} > 0$, the relationship between uncertainty and market rationality is no longer unambiguously positive as predicted by the canonical information theory. In particular, investor's signal precision responds to uncertainty as follows

$$\frac{\partial \tau_s^*}{\partial \tau_0^{-1}} = \tau_0^2 - \frac{r}{\lambda^2} \frac{\partial \lambda}{\partial \tau_0^{-1}}$$

So, when the information cost channel is strong enough, or when $\frac{\partial \lambda}{\partial \tau_0^{-1}}$ is sufficiently positive, the relation between uncertainty and investors' debiasing can turn negative, leading to a less rational market expectation.

### 3.2.4 Testable Predictions

We derive four testable predictions based on Proposition 1 and 2.

First, Proposition 1 predicts the heterogeneity in return predictability and the optimal attention's along the dimension of investors' financial stake.

**Prediction 1.** *The return predictability of EHB is decreasing in $r$ and investors' attention ($C^*$) is increasing in $r$ .*

*Proof.* See Appendix B.3. $\square$

Intuitively, when investors' financial stake is high, investors pay more attention to debias and obtain a more precise signal. When financial stakes relative to information cost are high enough, investors fully unravel analysts' bias, and there should be no return predictability.

Next, we test the implications of the presence of the information cost channel. First, building on Proposition 2, we test the implications when the strength of the information cost channel differs. Specifically, while both time-series and cross-sectional variations in uncertainty ($\tau_0^{-1}$) can be positively related to the unit information processing cost, intuitively, $\frac{\partial \lambda}{\partial \tau_0^{-1}}$ should be much more positive for cross-sectional than time-series variations in uncertainty. This is because a firm's information processing cost is determined by the information environment, which can vary significantly across firms but is unlikely to change rapidly over time. This leads us to the following prediction.

**Prediction 2.** *If the cross-sectional correlation between uncertainty and the information processing costs is more positive than the corresponding time-series correlation, the strength of* *return predictability of EHB can covary positively with cross-sectional variations in uncertainty but negatively with time-series uncertainty.*

Intuitively, when information costs do not vary significantly over short periods of time, sudden spikes in uncertainty could lead to an improvement in market rationality due to the information benefit channel. On the other hand, in the cross-section of firms where variations in firm-level uncertainty induce strong fluctuations in information processing costs, a negative relationship between uncertainty and market rationality may be observed.

Second, we directly test the role of information costs in shaping the heterogeneity in market rationality across firms and over time.

**Prediction 3.** *The return predictability of EHB is stronger among firms with high information processing costs; such return predictability should be weakened by information disclosure.*

*Proof.* See Appendix B.3. □

Finally, besides return predictability, the presence of the information cost channel also leads to distinct predictions for investors' attention allocation compared to the canonical information choice models without the information cost channel (i.e. $\frac{\partial \lambda}{\partial \tau_0^{-1}} = 0$).

**Prediction 4.** *Absent of the information cost channel ($\frac{\partial \lambda}{\partial \tau_0^{-1}} = 0$), both investors' attention and information acquisition respond positively to uncertainty, and a temporary increase in uncertainty leads to more attention allocation to firms with lower uncertainty. In the presence of the information choice channel ($\frac{\partial \lambda}{\partial \tau_0^{-1}} > 0$), investors' attention and information acquisition may respond to uncertainty in different ways and a temporary increase in uncertainty leads to more attention allocation to firms with higher uncertainty.*

*Proof.* See Appendix B.3. □

To test these predictions, we propose empirical measures and describe the data and the measure in the following section.

## 3.3 Data and Empirical Measurements

### 3.3.1 Data

Our sample consists of U.S. common stocks that are covered in the intersection of CRSP and Compustat. We exclude micro-cap stocks, defined as stocks with a market capitalization below the NYSE 20th percentile, and low price stocks, defined as stocks with a price below $5. We provide more details about the data and universe in Appendix B.1.

### 3.3.2 Empirical Measurements

We briefly describe the empirical measurements constructed for our analysis in this subsection. Please see Table 3.1 for a summary of all key variables in our analysis and see AppendixB.1 for a more detailed description of the variable construction.

#### Conditional Biases in Analysts' Earnings Forecasts

We quantify the rationality in market's earnings expectations using the return predictability of conditional biases in analysts' earnings forecasts. To construct the ex-ante measure of the conditional biases in analysts' earnings forecasts, we utilize machine learning algorithms.

Here we provide a brief description of how we construct the variable and leave more details in Appendix B.2.

The target variable is the realized errors of analysts' forecasts, defined as the median consensus EPS forecast minus the realized EPS.[13] We focus on predicting analysts' forecast errors associated with one-quarter- and two-quarter-ahead quarterly forecasts and one-year- and two-year-ahead annual forecasts (referred to as FQ1, FQ2, FY1, and FY2 hereafter).

Our predictor set consists of the financial ratios from the WRDS financial ratio suite, similar to de Silva and Thesmar, 2021; J. van Binsbergen et al., 2020, as well as the log of market capitalization, the current stock price, the stock return for the current month, and the realized earnings for latest fiscal year end and fiscal quarter end.[14] Analysts' forecasts contain private information that adds incremental predictive power for earnings relative to financial statement variables (de Silva and Thesmar, 2021; J. van Binsbergen et al., 2020). We thus also include the following analysts' forecasts related variables in our predictor set: median consensus forecasts for FQ1, FQ2, FY1, and FY2, the corresponding three-month revisions and the latest realized forecast errors of the these forecasts, and the number of months between the forecast date and the fiscal period end. The last variable captures the walk-down effect documented in Richardson et al., 2004.

[13]We use the unadjusted median consensus forecasts and the actual earnings from I/B/E/S. We account for the stock split effect by applying the cumulative adjustment factors from the CRSP, following Diether et al., 2002. In the earnings forecast literature, there are a variety of forecasting approaches. We review this literature in a companion paper. Our results are robust to reasonable changes in forecasting approaches, which we discuss in the Appendix B.2.

[14]We find qualitatively similar results when replacing the WRDS financial ratios with the predictors used in So, 2013 and major income statement and cash flow statement variables, or the predictors from Green et al., 2017, or the predictors from Chen and Zimmermann, 2022.

We use the gradient boosted decision tree model implemented under the LightGBM framework (LGBM), a popular, off-the-shelf machine learning algorithm, as our main statistical forecasting model.[15] Compared to linear models, such as those used in So, 2013, regression tree models are able to capture non-linear relationships in the data, which is important for forecasting future earnings, as shown in J. van Binsbergen et al., 2020.[16]

We train the regression tree model in rolling 10-year estimation windows and ensure our forecasts are strictly out-of-sample. We make sure that both the target variable and the predictors in the train dataset are known at month $t$. Specifically, the target variable in the train dataset is based on the realized earnings announced between $t - 120$ and $t$.[17] After we fit the model, which includes selecting the optimal hyper-parameters, we then apply the fitted model on the predictor values at the month $t$ to arrive at our ex-ante predictions of analysts' errors. We refer to this measure as the ex-ante human bias (EHB).

**Uncertainty Measures**

We employ six distinct monthly uncertainty measures to ensure our metrics cover various types of uncertainties, including cross-sectional versus time-series, firm-specific versus

[15]At a high level, these tree-based algorithms partition the data into groups that are similar in terms of sorting variables and then use the average value of the target variable within each partition to approximate the non-linear predictive relation. To grow a tree is to find the sorting variable from the predictor set and the split value to partition the data in a sequential way, such that the resulting approximation achieves the minimum forecast errors. See Campbell, Lu, Ham, and Wood (2023) for more detailed discussions of the machine learning earnings forecasting models.

[16]For robustness, we use alternative models, including the Random Forest model (J. H. van Binsbergen et al., 2022), Neural Network models and the linear models as in So, 2013 to perform our main tests regarding uncertainty and return predictability and find similar results. For more information on our implementation of these alternative algorithms, see Section B.2.1

[17]We follow Johnson and So, 2018 and construct the announcement dates of the earnings for a FPEDATS (forecast period end date) as the earlier one of the I/B/E/S announcement date (+1 if reported after 4pm) and Compustat announcement date. If neither IBES or Compustat announcement date is available, we assume the earnings are known six months after fiscal period end.

aggregate, and realized versus forward-looking uncertainties. Our main firm-level uncertainty measure is idiosyncratic volatility (IVOL), defined as the standard deviation of residuals from CAPM regressions using the past year of daily data, as defined in Ali et al., 2003. Since IVOL is based on realized volatility, we also use the option implied volatility (OIV) as a forward-looking firm-level uncertainty measure. OIV is the average implied volatility from a call and put option with 30 days to maturity and a delta of 0.5 (-0.5 for a put option) from the volatility surface file of OptionMetrics on the last day of the month. We further consider two measures related to IVOL: the 36 month rolling average ($\text{IVOL}_{MA36}$) and the ratio between IVOL and $\text{IVOL}_{MA36}$ (Abnormal IVOL). We use $\text{IVOL}_{MA36}$ to measure the long-run cross-sectional variation in firm-level uncertainty and Abnormal IVOL to measure the time-series variation in firm-level uncertainty.

We use VIX as our main measure for aggregate uncertainty. Since VIX captures the uncertainty reflected in the financial market, we also use the Economic Policy Uncertainty Measure (EPU) provided by Baker et al., 2016 as an alternative measure for aggregate uncertainty.[18]

We also use two daily measures of uncertainty: the daily measure of OIV as well as HLtH, which is a measure of intraday volatility. HLtH is the ratio between the stock's daily high and low price difference and the daily high price as described in Ben-Rephael et al., 2017. From these two measures, we also create abnormal OIV and HLtH measures which are the

[18]According to Baker et al., 2016, EPU "capture(s) uncertainty about who will make economic policy decisions, what economic policy actions will be undertaken and when, and the economic effects of policy actions (or inaction)."

ratio of the daily OIV or HLtH value to the average value of the measure from the previous 21 trading days (i.e. one month).

**Investor Attention**

Our main measure for investors attention is the abnormal investor attention (AIA) measure from Ben-Rephael et al., 2017.[19] As discussed in detail by Ben-Rephael et al., 2017, AIA is a direct measure of institutional investors' (money managers') attention because it directly measures the attention of Bloomberg terminal users through reading news articles about a firm or actively searching for information on a firm.[20] AIA is based on the daily Bloomberg News Heat-Daily Max Readership value (hereafter Bloomberg Attention), which has five integer values ranging from 0 to 4 that captures the attention relative to the previous 30 days. AIA is an indicator variable equal to one for days with a Bloomberg Attention measure of 3 or 4. To create a monthly measure, we take the average of the daily measure across a given month.

## 3.4 Empirical Results

We present four sets of empirical results corresponding to the four predictions in Section 3.2.4.

[19] See here for their detailed methodology: https://www3.nd.edu/∼zda/

[20] Actively searching for a firm has a greater weight in the creation of the attention measure relative to a terminal user clicking on a news article.

### 3.4.1 The Relationships Between Market Rationality and Financial Stake

Prediction 1 posits that a firm's market capitalization ($r$) is an important factor in driving return predictability of EHB and investors' attention. We first conduct portfolio sort analysis to examine the relation between firms' market capitalization and return predictability of EHB. Specifically, at the end of each month, we sort firms into three groups based on their market capitalization using the 20th, 50th, and 80th percentiles of the size distribution of NYSE stocks. Small-cap firms are those between the 20th and 50th percentiles, mid-cap firms are those between the 50th and 80th percentiles, while mega-cap firms refer to those larger than the 80th percentile.[21]

Conditional on firms' size group, we further sort firms equally into quintiles based on the ascending order of EHB. Finally, we calculate the value-weighted return over the next month for each of the 15 Size-EHB portfolios.

Panel A of Table 3.2 presents portfolio excess returns. The EHB Q1-Q5 portfolio generates a statistically significant excess return spread of 0.743% per month among small-cap stocks; whereas for mid-cap and large-cap stocks, the excess return spreads are not significant. As a result, the difference between the long-short EHB portfolios of the small-cap vs. mega-cap has a significant return difference of 0.564% per month.

[21]After dividing firms into these size groups, we have around 19% of firms in the Mega-Cap group, 33% of firms in the mid-cap group, and 48% of firms in the small cap group. Note that our sample already excludes micro-cap stocks, which are below the NYSE 20th percentile.

To validate whether the return spreads between portfolios are not due to their systematic loading on risk factors, Panel B of Table 3.2 presents results of portfolio FF5 alphas. The EHB Q1-Q5 for small-cap firms have an FF-5 alphas of 0.937% per month, higher than the excess returns. The increase in magnitude is consistent with the hypothesis that EHB is capturing the irrationality of the market's cash flow expectations, and therefore after controlling for systematic risk factors the effects should be stronger. Although the mid-cap firms also have a significant FF-5 alpha now, the magnitude is smaller than that of the small-cap, at 0.546% per month. The mega-cap EHB Q1-Q5 portfolio does not generate a statistically significant alpha. The small-cap EHB Q1-Q5 portfolio generates an alpha that is 0.634% per month larger than that of the mega-cap portfolio with a t-statistic of 2.88. This is consistent with the Prediction 1 in which the return predictability of EHB is decreasing as firm size increases.

The results based on investor attention AIA further confirm the model prediction. Specifically, we regress our investor attention measure AIA on Size. As AIA is an indicator variable, we use probit regressions with standard errors clustered by firm and day. Panel C of Table 3.2 show the regression coefficients for the different market segments (small-cap, mid-cap, and mega-cap) as well as the full sample. The estimated coefficients on firm size are positive and strongly significant. This results suggest investors are more likely to pay more attention to larger firms than to smaller firms, confirming Prediction 4 that investors allocate more of their attention when their financial stakes are higher.

Besides testing the model's prediction, the results in Table 3.2 provide a new perspective to understand the degree of rationality of the market's earnings expectation overall. J. H. van Binsbergen et al., 2022 propose the machine-learning based model to build earnings' expectation as a benchmark, and show that the analysts' ex-ante biases strongly predicts future returns, on average. The results in Table 3.2 show that this return predictability has large heterogeneity among firms of different sizes. In fact, as this table quantify , the EHB only predicts future returns among the 20% of the total market capitalization, which means that the market is mostly rational, or 80% rational, in terms of market capitalization.

In sum, the results presented here establish that a firm's size is an important factor driving the rationality of market expectation. Therefore, we control for firm size in the subsequent analysis whenever we can.

### 3.4.2 The Relationships Between Uncertainty and Market Rationality

Prediction 2 states that if the cross-sectional correlation between uncertainty and the information processing costs is more positive than the corresponding time-series correlation, the strength of return predictability of EHB can covary positively with cross-sectional variations in uncertainty but negatively with time-series uncertainty.

We test Prediction 2 using portfolio sorts. To account for the size effect documented in Section 3.4.1, we first divide firms into non-mega cap (small- and mid-cap) and mega-cap groups. Within each size group, we create double-sort portfolios based on measures of uncertainty and EHB. Specifically, for the uncertainty measures with cross-sectional variation (all but VIX and EPU), we begin by cross-sectionally sorting firms into terciles based on the specific firm-level uncertainty measure. For VIX and EPU, we split firms into terciles based on their historical values our sample period. T1 of each uncertainty measure contains firms (time periods) with the lowest values of uncertainty and T3 contains firms (time periods) with the highest values of uncertainty. Within each month, firms are then conditionally sorted into quintiles based on EHB (i.e. monthly rebalancing). EHB Q1 contains firms with the lowest values of EHB and EHB Q5 contains firms with the highest values. We finally calculate the value-weighted next-month return for each of the 15 portfolios.

**Cross-Sectional vs. Time-Series Uncertainty** We start with idiosyncratic volatility (IVOL) as our main cross-sectional measure of uncertainty and VIX as time-series measure of uncertainty. Table 3.3 presents the FF-5 alphas for the 15 portfolios based on IVOL and EHB, for non-mega and mega-cap stocks. We see that for firms with low uncertainty (IVOL T1), the portfolios never generate a statistically significant FF5 alpha across EHB quintiles. The FF-5 alphas of long-short portfolio (EHB Q1-Q5) is also insignificantly different from zero in both the non-mega and mega-cap firms.

However, as firms' uncertainty (IVOL) increases, the FF-5 alphas of EHB Q1-Q5 portfolio increase in magnitude as well as statistical significance. In the non-mega-cap firms, we see that the long-short portfolios of firms with higher IVOL,T2 and T3, generate statistically significant alphas of 0.529% and 1.234% per month, respectively. Consistent with financial stakes as an important factor driving market rationality, we see in the mega-cap groups, only the T3 (firms with the highest IVOL) long-short portfolio generates a statistically significant alpha of 0.871% per month. As a result, in both the non-mega-cap and mega-cap firms, we see that the firms with the greatest uncertainty (T3) generate 1.200% and 1.073% greater monthly alphas than firms with lower uncertainty, respectively. This supports Prediction 2.

As a summary, results in this table show greater return predictability of EHB for firms with greater uncertainty relative to others in the cross-section (i.e. those with higher information processing costs). This result can not be explained by canonical information choice models and supports the presence of a strong information cost channel in the cross-section.

We find the opposite results when analyzing the relationship between return predictability of EHB and VIX, our primary measure of time-series uncertainty. Table 3.4 shows the FF-5 alphas for the portfolios. We find positive and statistically significant alphas in periods of low uncertainty, especially in the non-mega cap firms. For the non-mega cap firms, periods of lowest uncertainty (T1) are associated with a statistically significant alpha of 0.659% per month whereas the L-S portfolio for the periods of greatest uncertainty (T3) fail to generate statistically significant alphas.

These contrasting relationships between uncertainty and return predictability of EHB are hard to reconcile using existing information choice models. On the other hand, they are consistent with a model with the information cost channel. To ensure robustness of our finding, we consider forward-looking measure of firm-level uncertainty next and show the results are robust whether we use IVOL or option-implied volatility as firm-level measure of uncertainty.

**Realized vs. Implied Volatility** We consider our second measure of cross-sectional uncertainty. Table 3.5 presents the results for the portfolio analysis using OIV as the measure of firm-level uncertainty. We find that our results are consistent with those using IVOL as a measure of cross-sectional uncertainty.

For firms with the lowest uncertainty (T1), the EHB Q1-Q5 long-short portfolio does not generate statistically significant alphas. However, for both the non-mega-cap and mega-cap firms, long-short portfolio containing the firms with the highest uncertainty (T3) generate statistically significant alphas of 1.176% and 0.966% per month, respectively. The firms in with the greatest uncertainty generate 1.019% (1.059%) per month greater alphas than the lowest uncertainty firms for the non-mega-cap (mega-cap) firms. This suggests that our findings to support Prediction 2 are not driven by realized volatility as using implied volatility as a measure of uncertainty generates similar findings.

**Fundamental vs. Financial Volatility** We further consider alternative measure of time-series uncertainty. Since VIX can spike due to temporary investor sentiment shocks other than fundamental uncertainty, we repeat the analysis using EPU which captures economic uncertainty. As with the IVOL and OIV results, we find similar patterns in Table 3.6 as were found in 3.4. As with the analysis using VIX, we show that for periods of heightened uncertainty (T3), our long-short EHB Q1-Q5 portfolios generate no statistically significant alphas in the non-mega-cap or mega-cap firms. For the periods of lower uncertainty (T1 and T2) the long-short portfolios generate statistically significant alphas for both the non-mega-cap and mega-cap firms, though the magnitude and significance is reduced for the mega-cap firms. In the non-mega-cap firms, the periods of low uncertainty generate 1.297% per month greater alphas than the periods of high uncertainty.

In sum, this result provides further support that the findings that support Prediction 2 are not driven by our selection of financial vs. economic measures of time-series uncertainty.

**Short-Run vs. Long-Run Volatility** IVOL and OIV capture more recent measures of uncertainty for the firms. However, one may argue firm-level fundamental uncertainty as well as the strength of the information cost channel should reflect more persistent firm-level fundamental uncertainty.

To ensure our results are robust to this concern, we use IVOL$_{MA36}$ as a proxy for long-run uncertainty as it is the moving average of IVOL over the previous 36 months. Table 3.7 shows these results. Again, we find that the long-short portfolio for firms with the lowest long-run uncertainty generate no statistically significant alphas. However, in the non-mega-cap firms, the T2 and T3 EHB Q1-Q5 portfolios generate statistically significant alphas of 0.505% and 0.998% per month, respectively. We also continue to find that the firms with the greatest uncertainty generate alphas that are statistically significantly greater than those of low uncertainty. For the non-mega-cap firms, this difference is 1.120% per month.

**Individual vs. Aggregate Volatility** One may argue the contrasting relationship between uncertainty and return predictability comes from the fact that our cross-sectional tests are on the firm-level while the time-series tests are on the aggregate-level. To ensure our results are not driven by this difference we conduct both the cross-sectional and time-series tests based on firm-level measures. Specifically, we use abnormal IVOL to capture time-series variation in uncertainty at the firm level.

Table 3.8 shows that our results are not due to the difference between aggregate vs. firm-level uncertainty measure. Indeed, as results in this table show, The EHB Q1-Q5 long-short portfolio for firms experiencing periods of high uncertainty (T3) generate no alphas. However, firms in periods of low uncertainty (T1) generate statistically significant alphas for both the non-mega-cap and mega-cap firms of 0.710% and 0.476% per month, respectively.

In sum, we have shown that the contrasting relationships between uncertainty and market rationality are a robust phenomenon in the data. These relationships are consistent with the presence of a information cost channel within the information choice model, consistent with the Prediction 2. To further test the presence of the information cost channel, we present results on Prediction 3 next.

### 3.4.3 Information Cost, Market Rationality and Earnings Announcements

Prediction 3 suggests that for firms with high information processing costs have stronger return predictability of EHB. We use firm age as our proxy for information cost processing. Younger firms naturally have less information available for investors, making it more challenging and costly to unravel the analysts' biases. Specifically, we set our analysis up using 1/Age in situations where we use portfolio sorts so firms with the highest processing costs (young firms) are in the highest tercile (T3).

Table 3.9 shows the portfolio sort analysis for variations in information costs. As with the uncertainty measures, we first divide firms into non-mega cap (small- and mid-cap) and mega-cap groups as described in Section 3.4.1. We then cross-sectionally sort firms into terciles based on 1/Age. Within each month, firms are then conditionally sorted into quintiles based on EHB (i.e. monthly rebalancing). EHB Q1 contains firms with the lowest values of EHB and EHB Q5 contains firms with the highest values. We finally calculate the value-weighted next-month return for each of the 15 portfolios.

For firms with low information processing costs (T1), the long-short portfolio generates no statistically significant alphas. However, as information processing costs increase, we find that in the non-mega-cap firms, the higher information processing cost portfolios, T2 and T3, generate statistically significant alphas of 0.664% and 1.274% per month respectively. The firms with the highest processing costs for the non-mega-cap firms generate alphas that are 1.170% per month greater than those with low processing costs. In the mega-cap firms, only the T3 firms generate alphas of 0.655% per month. This provides support for Prediction 3 as we find that young firms, which have higher information processing costs have a stronger return predictability of EHB as compared to firms with lower information processing costs.

Prediction 3 also suggests that the return predictability results should weaken as new information is disclosed. To test for this, we evaluate differences between months with and without earnings announcements for firms. Earnings announcements are one of the largest sources of information disclosure for a firm, so we expect the return predictability results to be reduced for firms in months with earnings announcements as compared to non-earnings announcements months. To do so, we use Fama MacBeth regressions.

We first show that the results found in Table 3.9 are robust using a Fama MacBeth regression setup. Table 3.10 cross-sectionally sorts firms into terciles based on 1/Age, so once again, firms with the highest information processing costs are in T3. We show that the return predictability of EHB is statistically significant in the high information processing cost firms but is not statistically significant in the low information processing costs firms. This is robust to a univariate regression or the inclusion of controls.

We then repeat the analysis by dividing firms based on earnings announcement months. Table 3.11 shows the results for non-earnings announcement firm-months in Panel A, earnings announcement firm-months in Panel B and the difference between the two groups in Panel C. We find that, in Panel A, firms in months without earnings announcements show the same pattern as the full sample. Firms with high information processing costs have positive and statistically significant return predictability. The difference between firms with high and low information processing costs is statistically significant and robust to controls. However, in months with earnings announcements, we find that the magnitude and significance of our results are reduced in the univariate setting. When including controls, the return predictability of EHB becomes insignificant for the high information cost firms. We also find no statistically significant difference between the high and low information processing cost firms. This supports Prediction 3 that when new information is disclosed, the return predictability of EHB weakens. Panel C further provides support to Prediction 3. We show that the difference in the return predictability of EHB between earnings announcement months and non-earnings announcement months is statistically significant for firms with higher information processing costs (T2 and T3).

### 3.4.4 The Cross-Section and Time-Series of Attention Allocation Facing Uncertainty

Our final analysis is for Prediction 4. This predicts that absent of the information cost channel, as uncertainty increases, investors allocate more attention to firms. We analyze the impact of uncertainty on investors' attention in Table 3.12. The results show that as firm uncertainty increases, investors' attention increases. We see this hold for the daily OIV and HLtH uncertainty measures as well as the abnormal measures that capture spikes in firm uncertainty (Abnormal OIV and Abnormal HLtH). These results suggest that investors' are more likely to pay attention to firms as uncertainty increases.

Additionally, absent of the information cost channel, investors are expected to allocate more attention to firms with lower long-run uncertainty in periods where these firms experience spikes in uncertainty. In the presence of the information cost channel, our model predicts that investors will allocate greater attention to firms with higher long-run uncertainty when these firms experience spikes in uncertainty.

We investigate how investors' allocate attention when firms experience spikes in uncertainty and compare firms with high and low long-run uncertainty. To do so we first divide firms into terciles based on $\text{IVOL}_{MA36}$ orthogonalized to Size. This captures the long-run firm uncertainty after controlling for the effects of $r$. As our attention results are on a daily basis, we use the $\text{IVOL}_{MA36}$ measure from the end of the previous month to sort firms cross-sectionally into terciles. We use the two daily abnormal uncertainty measures to capture spikes in firm uncertainty (Abnormal OIV and Abnormal HLtH) and evaluate the relationship with investors' attention.

Table 3.13 show the results of the analysis. We show that overall, spikes in uncertainty lead to an increase in investor attention for both high and low long-run uncertainty firms. However, when we consider the difference in the magnitude of the the likelihood of investors' increasing their attention for high and low long-run uncertainty firms, we see that firms with higher long-run uncertainty experience larger increases in investors' attention when there are spikes in uncertainty. We show that the increase in attention for high long-run uncertainty firms is statistically significantly larger than for low long-run uncertainty firms. This result supports the information cost channel as if the information cost channel were absent, we would expect the attention to increase more for low long-run uncertainty firms than for high long-run uncertainty firms.

## 3.5 Reconciling Existing Findings

Zhang, 2006 shows that analysts' forecast revisions predict stock returns more positively when uncertainty is high. Zhang, 2006's finding is commonly interpreted as suggesting investors react less completely to analysts' forecast revisions when uncertainty is high.[22] However, this under-reaction mechanism appears inconsistent with the key finding in Loh and Stulz, 2018 that the price impact of analysts' forecast revisions is higher when (aggregate

[22] Zhang, 2006 is agnostic about why investors would have differential (under-) reactions, as he notes "Because I do not incorporate measures of private information or overconfidence in my empirical analysis, my evidence leaves the door open for other behavioral models. For example, my results are also consistent with a behavioral model in which investors overweight their priors relative to new information due to the anchoring/conservatism bias and over weight their priors more when there is greater information uncertainty".

or firm-level) uncertainty is high, which suggests investors react more strongly to analysts' forecasts when facing more uncertainty.

Our model reconciles these seemly contradictory findings. In our model, investors endogenously choose to place a higher weight on analysts' forecasts when facing increased uncertainty, which explains why the price impact of analysts' forecast revisions is larger when uncertainty is high. However, our model generates Zhang, 2006's finding of a positive cross-sectional relation between uncertainty and post-revision drift. This is not because investors underreact more to the information in analysts' forecasts but rather because investors overreact more to the bias in analysts' forecasts (i.e., debias less) among stocks with high uncertainty.

Specifically, it is well known that analysts' forecasts become increasingly less biased as earnings announcements approach, so analysts' forecasts that have a positive (negative) EHB will be revised down (up) gradually, leading to a negative (positive) revision. Since EHB negatively predict returns most strongly among high uncertainty firms, if EHB and revision are negatively correlated, then revisions will positively predict returns most strongly among high uncertainty firms.

To empirically evaluate the merit of our explanation, we first compute the Spearman rank correlation between EHB and revisions in Table 3.14. Consistent with our hypothesis, we find that analysts' revisions and EHB are highly negatively correlated, ranging from -0.569 to -0.608 across the IVOL terciles.

To more directly test our hypothesis that the return predictability of EHB is driving Zhang, 2006's finding, we run a direct horse race between EBH and analysts' forecast revisions. In Table 3.15, we follow the regression specification in Zhang, 2006 and run a Fama MacBeth regression of the next-month returns on a revision dummy, which is equal to 1 if the 3-month revision (average over FQ1/FQ2/FY1/FY2) is positive, -1 if negative, and 0 otherwise. We replicate Zhang's result in our sample that the post-revision drift is larger among the top IVOL tercile compared to the bottom IVOL tercile, and it is only significant among the former. We then include EHB in the regressions. We find that the revision dummy is no longer significant when we control for EHB in the top IVOL tercile, whereas the coefficient on EHB is negative and significant. We repeat these analysis in the full sample and find a similar pattern that EHB subsumes the return predictive power of the revision dummy.

## 3.6 Conclusion

We document that time-series variations in uncertainty and cross-sectional variations in uncertainty are related to the return predictability of analysts' biases in opposite ways. These results are important because they contrast predictions of existing theories.

We propose an information choice model to understand these contrasting relationships. In our model, investors gain a more precise signal about future earnings if they unravel the conditional biases in analysts' forecasts, but they must acquire costly information in order to do so. Uncertainty is related to the investors' information choice and the resulting rationality of market expectations through two channels. First, through the information benefit channel, higher uncertainty means an additional unit of information resolves more uncertainty, and thus investors acquire more information to debias analysts' forecasts. Second, through the information cost channel, higher uncertainty indicates that less information is available, and thus information costs are high. Investors optimally choose to acquire less information to debias analysts' forecasts.

Our model provides concrete mechanisms through which uncertainty affects investors' expectations, which leads to novel implications regarding the relationship not only between uncertainty and the return predictability of analysts' conditional biases, but also between uncertainty and investors' attention allocation. Consistent with our hypothesis that the cross-sectional relationship between uncertainty and the return predictability of EHB is driven by the information cost channel, we find that variations in information costs, as measured by 1/Age and the residual financial news coverage, are positively correlated with firm-level uncertainty, negatively correlated with investors' attention, and positively correlated with the return predictability of EHB. Consistent with our hypothesis that the time-series relationship between uncertainty and the return predictability of EHB is driven by the information benefit channel, we find that investors' attention is high and the return predictability of EHB is weak when the aggregate uncertainty – proxied by VIX and EPU – is high or when the within-firm variation in uncertainty – proxied by earnings announcements and abnormal IVOl – is high.

Finally, our model shows that when uncertainty is high, investors can put a large weight on analysts' forecasts while debiasing less. This explains why when uncertainty is high, analysts' revisions can lead to both a greater contemporaneous price changes (Loh and Stulz, 2018) and a greater post-revision price drift (Zhang, 2006). Overall, our model not only explains why investors do not fully unravel analysts' biases decades after the return predictability of these biases is known, but also explains why different variations in uncertainty can affect the deviation of market earnings expectations from FIRE in opposite ways.

Table 3.1: Key Variable Definitions

This table provides the definition of key variables in the analysis.

| Variable | Definition |
|---|---|
| Ex-Ante Human Bias (EHB) | Analysts' conditional biases (Average across FQ1/FQ2/FY1/FY2 forecast horizons) |
| Age | Firm Age (months) |
| Size | Ln(Market Capitalization) (daily or as of end of month) |
| Mega Cap | Firms with market capitalization above 80th percentile of NYSE firm size |
| Small Cap | Firms below median NYSE market capitalization |
| Mid Cap | Firms above median NYSE market capitalization but not Mega Cap |
| IVOL | Standard deviation of residuals from CAPM regressions using the past year of daily data. (Require at least 100 non-missing observations.) |
| OIV | Average of call and put option implied volatility from the volatility surface using 30 day maturity and delta=0.5 (-0.5 for put options) on last day of the month. |
| $\text{OIV}_{MA21}$ | Moving average of OIV from day t-22 to t-1 |
| Abnormal OIV | $\frac{OIV}{OIV_{MA21}}$ |
| $\text{IVOL}_{MA36}$ | Moving Average of IVOL from month $t-35$ to $t$ (Trailing IVOL) |
| Abnormal IVOL | $\frac{IVOL}{IVOL_{MA36}}$ |
| VIX | VIX on last day of a given month |
| EPU | Economic Policy Uncertainty Index (Baker et al. 2014) |
| HLtH | The ratio between the stock's daily high and low price difference and the daily high price (Ben Rephael et al. 2017) |
| $\text{HLtH}_{MA21}$ | Moving average of HLtH from day t-22 to t-1 |
| Abnormal HLtH | $\frac{HLtH}{HLtH_{MA21}}$ |
| Bloomberg Attention Score | Monthly Average of Bloomberg News Heat-Daily Max Readership Measure (Values range from 0-4) |
| AIA | Indicator Variable equal to 1 when Bloomberg Attention score is 3-4 and 0 otherwise |

Table 3.2: Firm Size, the Return Predictability of Analysts' Conditional Biases, and Investors Attention

This table presents the analysis of size and the return predictability of analysts' conditional biases and size and investors' attention. Panels A and B show the excess returns and Fama French Five Factor alphas for double sort portfolios created by cross-sectionally sorting companies into groups based on their market capitalization then conditionally sorting firms into quintiles based on their EHB. The market value of equity groups divide the firms into mega-cap, mid-cap, and small cap groups. Mega-cap firms are defined as firms with market capitalization greater than the 80th percentile of firm sizes on the NYSE. The remaining firms are then defined as small- or mid-cap based on whether their size is above the median NYSE market capitalization. Panel A shows the excess returns and panel B shows Fama French Five factor abnormal return in percentage terms. Portfolios are value weighted and are re-balanced on a monthly basis. Standard errors of the resulting regression coefficients are computed based on Newey and West, 1987 with 12 lags. Q1 indicates the lowest values and Q5 the highest values for EHB. Panel C uses probit regression to analyze the relationship between AIA and Size. Firms are divided into into mega-cap, mid-cap, and small cap groups before running the probit regression of AIA on Size. For days when Bloomberg has no value, we do not fill the missing information of AIA with zero. Standard errors are clustered by firm and day. Statistical significance is denoted as ***, **, and * for $p<0.10$, $p<0.05$, and $p<0.01$, respectively. The sample period is from June 1990 to December 2019 for Panel A and B. The sample period for Panel C is from March 2010 to December 2019.

Panel A: Excess Return

| | EHB Q1 | EHB Q2 | EHB Q3 | EHB Q4 | EHB Q5 | EHB Q1-Q5 |
|---|---|---|---|---|---|---|
| Small Cap | 1.215*** | 0.906*** | 0.774*** | 0.690** | 0.472 | 0.743*** |
| | (4.73) | (3.80) | (2.91) | (2.34) | (1.21) | (2.96) |
| Mid Cap | 0.963*** | 0.828*** | 0.770*** | 0.759*** | 0.654* | 0.309 |
| | (3.83) | (3.73) | (3.41) | (2.97) | (1.82) | (1.40) |
| Mega Cap | 0.769*** | 0.685*** | 0.595*** | 0.750*** | 0.589* | 0.179 |
| | (2.63) | (2.82) | (2.86) | (3.16) | (1.86) | (0.86) |
| Small-Mega | 0.446* | 0.221 | 0.179 | -0.060 | -0.117 | 0.564** |
| | (1.93) | (1.21) | (0.98) | (-0.34) | (-0.63) | (2.15) |

Panel B: Abnormal Return

| | EHB Q1 | EHB Q2 | EHB Q3 | EHB Q4 | EHB Q5 | EHB Q1-Q5 |
|---|---|---|---|---|---|---|
| Small Cap | 0.368*** | 0.147*** | -0.029 | -0.159 | -0.568*** | 0.937*** |
| | (3.96) | (2.79) | (-0.50) | (-1.57) | (-3.15) | (3.74) |
| Mid Cap | 0.255** | 0.193*** | 0.008 | -0.108 | -0.292* | 0.546** |
| | (2.12) | (2.61) | (0.11) | (-1.23) | (-1.71) | (2.14) |
| Mega Cap | 0.193* | 0.044 | -0.060 | 0.068 | -0.109 | 0.302 |
| | (1.77) | (0.59) | (-0.95) | (0.91) | (-0.89) | (1.48) |
| Small-Mega | 0.175 | 0.103 | 0.030 | -0.226** | -0.459*** | 0.634*** |
| | (1.52) | (1.38) | (0.33) | (-2.47) | (-3.00) | (2.88) |

Panel C: Firm Size and Investors' Attention

| | Small Cap | Mid Cap | Mega Cap | Full Sample |
|---|---|---|---|---|
| Size | 0.055*** | 0.050*** | 0.206*** | 0.173*** |
| | (4.1) | (3.5) | (15.7) | (33.7) |
| Observations | 932661 | 957288 | 703925 | 2593874 |
| Pseudo $R^2$ | 0.000 | 0.000 | 0.017 | 0.031 |

Table 3.3: IVOL and the Return Predictability of Analysts' Conditional Biases

This table presents the Fama French Five Factor alphas for double sort portfolios created by cross-sectionally sorting companies into terciles based on IVOL then conditionally cross-sectionally sorting firms into quintiles based on their EHB. Firms are split into mega- and non-mega-cap firms prior to creating the portfolios. Panel A shows the results for non-mega-cap firms and panel B shows the results for mega-cap firms. Portfolios are value weighted and are re-balanced on a monthly basis. Standard errors of the resulting regression coefficients are computed based on Newey and West, 1987 with 12 lags. Statistical significance is denoted as ***, **, and * for $p<0.10$, $p<0.05$, and $p<0.01$, respectively. T1 indicates the lowest values and T3 the highest values for IVOL. Q1 indicates the lowest values and Q5 the highest values for EHB. The sample period is from June 1990 to December 2019.

Panel A: Non-Mega Cap

| | EHB Q1 | EHB Q2 | EHB Q3 | EHB Q4 | EHB Q5 | EHB Q1-Q5 |
|---|---|---|---|---|---|---|
| IVOL T1 | -0.003 | 0.092 | 0.031 | -0.009 | -0.037 | 0.034 |
| | (-0.03) | (1.00) | (0.40) | (-0.10) | (-0.28) | (0.21) |
| IVOL T2 | 0.196 | 0.012 | -0.183** | -0.140 | -0.333* | 0.529** |
| | (1.61) | (0.12) | (-1.99) | (-1.24) | (-1.84) | (2.20) |
| IVOL T3 | 0.598*** | 0.198 | -0.159 | -0.605*** | -0.636*** | 1.234*** |
| | (2.60) | (1.30) | (-1.32) | (-3.60) | (-2.59) | (3.19) |
| IVOL T1-T3 | -0.601** | -0.105 | 0.190 | 0.596*** | 0.599*** | -1.200*** |
| | (-2.14) | (-0.48) | (1.14) | (3.85) | (2.70) | (-3.55) |

Panel B: Mega Cap

| | EHB Q1 | EHB Q2 | EHB Q3 | EHB Q4 | EHB Q5 | EHB Q1-Q5 |
|---|---|---|---|---|---|---|
| IVOL T1 | -0.086 | 0.059 | -0.108 | 0.045 | 0.116 | -0.202 |
| | (-0.76) | (0.71) | (-1.18) | (0.51) | (1.10) | (-1.22) |
| IVOL T2 | 0.185 | 0.007 | -0.078 | 0.043 | -0.167 | 0.352 |
| | (1.30) | (0.06) | (-0.88) | (0.35) | (-1.31) | (1.54) |
| IVOL T3 | 0.568** | 0.258 | -0.063 | 0.204 | -0.302 | 0.871** |
| | (2.24) | (1.60) | (-0.44) | (1.20) | (-1.45) | (2.50) |
| IVOL T1-T3 | -0.654** | -0.198 | -0.044 | -0.159 | 0.418* | -1.073*** |
| | (-2.47) | (-1.16) | (-0.24) | (-0.88) | (1.76) | (-2.84) |

Table 3.4: VIX and the Return Predictability of Analysts' Conditional Biases

This table presents the Fama French Five Factor alphas for double sort portfolios created by sorting companies into terciles based on VIX then conditionally cross-sectionally sorting firms into quintiles based on their EHB. Firms are split into mega- and non-mega-cap firms prior to creating the portfolios. Panel A shows the results for non-mega-cap firms and panel B shows the results for mega-cap firms. Portfolios are value weighted and are re-balanced on a monthly basis. Standard errors of the resulting regression coefficients are computed based on Newey and West, 1987 with 12 lags. Statistical significance is denoted as ***, **, and * for $p<0.10$, $p<0.05$, and $p<0.01$, respectively. T1 indicates the lowest values and T3 the highest values for VIX. Q1 indicates the lowest values and Q5 the highest values for EHB. The sample period is from June 1990 to December 2019.

Panel A: Non-Mega Cap

| | EHB Q1 | EHB Q2 | EHB Q3 | EHB Q4 | EHB Q5 | EHB Q1-Q5 |
|---|---|---|---|---|---|---|
| VIX T1 | 0.231** | 0.167** | -0.005 | -0.080 | -0.428*** | 0.659*** |
| | (2.20) | (2.56) | (-0.06) | (-0.83) | (-2.65) | (2.80) |
| VIX T2 | 0.253*** | 0.279*** | -0.170* | -0.311** | -0.435** | 0.689*** |
| | (2.87) | (2.93) | (-1.71) | (-2.37) | (-2.50) | (3.23) |
| VIX T3 | 0.165 | 0.002 | 0.043 | 0.099 | 0.104 | 0.061 |
| | (0.84) | (0.02) | (0.35) | (0.55) | (0.28) | (0.12) |
| VIX T1-T3 | 0.066 | 0.165 | -0.048 | -0.179 | -0.532 | 0.597 |
| | (0.29) | (1.13) | (-0.32) | (-0.88) | (-1.32) | (1.07) |

Panel B: Mega Cap

| | EHB Q1 | EHB Q2 | EHB Q3 | EHB Q4 | EHB Q5 | EHB Q1-Q5 |
|---|---|---|---|---|---|---|
| VIX T1 | 0.083 | 0.126 | -0.225** | 0.115 | -0.173 | 0.256 |
| | (0.61) | (1.08) | (-2.42) | (1.31) | (-1.03) | (0.99) |
| VIX T2 | 0.293** | 0.118 | -0.019 | -0.036 | -0.355** | 0.648** |
| | (2.09) | (1.10) | (-0.15) | (-0.38) | (-2.19) | (2.45) |
| VIX T3 | 0.095 | -0.068 | -0.014 | 0.031 | 0.260 | -0.165 |
| | (0.46) | (-0.48) | (-0.13) | (0.28) | (0.95) | (-0.38) |
| VIX T1-T3 | -0.012 | 0.194 | -0.212 | 0.084 | -0.433 | 0.421 |
| | (-0.05) | (1.06) | (-1.50) | (0.58) | (-1.35) | (0.85) |

Table 3.5: OIV and the Return Predictability of Analysts' Conditional Biases

This table presents the Fama French Five Factor alphas for double sort portfolios created by cross-sectionally sorting companies into terciles based on OIV then conditionally cross-sectionally sorting firms into quintiles based on their EHB. Firms are split into mega- and non-mega-cap firms prior to creating the portfolios. Panel A shows the results for non-mega-cap firms and panel B shows the results for mega-cap firms. Portfolios are value weighted and are re-balanced on a monthly basis. Standard errors of the resulting regression coefficients are computed based on Newey and West, 1987 with 12 lags. Statistical significance is denoted as ***, **, and * for $p<0.10$, $p<0.05$, and $p<0.01$, respectively. T1 indicates the lowest values and T3 the highest values for OIV. Q1 indicates the lowest values and Q5 the highest values for EHB. The sample period is from January 1996 to December 2019.

Panel A: Non-Mega Cap

| | EHB Q1 | EHB Q2 | EHB Q3 | EHB Q4 | EHB Q5 | EHB Q1-Q5 |
|---|---|---|---|---|---|---|
| OIV T1 | -0.003 | 0.119 | -0.050 | -0.080 | -0.160 | 0.157 |
| | (-0.03) | (1.16) | (-0.51) | (-0.77) | (-1.08) | (0.93) |
| OIV T2 | 0.240* | 0.041 | -0.082 | -0.061 | -0.148 | 0.389 |
| | (1.79) | (0.51) | (-0.72) | (-0.47) | (-0.61) | (1.20) |
| OIV T3 | 0.237 | 0.251 | -0.150 | -0.607** | -0.938*** | 1.176*** |
| | (1.11) | (1.56) | (-0.85) | (-2.20) | (-3.09) | (2.78) |
| OIV T1-T3 | -0.240 | -0.132 | 0.100 | 0.527** | 0.779*** | -1.019*** |
| | (-0.98) | (-0.54) | (0.53) | (2.08) | (2.90) | (-2.86) |

Panel B: Mega Cap

| | EHB Q1 | EHB Q2 | EHB Q3 | EHB Q4 | EHB Q5 | EHB Q1-Q5 |
|---|---|---|---|---|---|---|
| OIV T1 | -0.114 | -0.027 | -0.145 | -0.068 | -0.021 | -0.093 |
| | (-0.85) | (-0.25) | (-1.51) | (-0.64) | (-0.15) | (-0.51) |
| OIV T2 | 0.233* | 0.107 | -0.005 | -0.093 | -0.139 | 0.372* |
| | (1.74) | (1.09) | (-0.05) | (-0.63) | (-1.01) | (1.78) |
| OIV T3 | 0.433 | 0.560*** | 0.121 | 0.270 | -0.533** | 0.966** |
| | (1.28) | (2.77) | (0.69) | (1.39) | (-2.36) | (2.06) |
| OIV T1-T3 | -0.547 | -0.587** | -0.266 | -0.339 | 0.512* | -1.059* |
| | (-1.35) | (-2.57) | (-1.39) | (-1.37) | (1.68) | (-1.87) |

Table 3.6: EPU and the Return Predictability of Analysts' Conditional Biases

This table presents the Fama French Five Factor alphas for double sort portfolios created by sorting companies into terciles based on EPU then conditionally cross-sectionally sorting firms into quintiles based on their EHB. Firms are split into mega- and non-mega-cap firms prior to creating the portfolios. Panel A shows the results for non-mega-cap firms and panel B shows the results for mega-cap firms. Standard errors of the resulting regression coefficients are computed based on Newey and West, 1987 with 12 lags. Statistical significance is denoted as ***, **, and * for $p<0.10$, $p<0.05$, and $p<0.01$, respectively. T1 indicates the lowest values and T3 the highest values for EPU. Q1 indicates the lowest values and Q5 the highest values for EHB. The sample period is from June 1990 to December 2019.

Panel A: Non-Mega Cap

| | EHB Q1 | EHB Q2 | EHB Q3 | EHB Q4 | EHB Q5 | EHB Q1-Q5 |
|---|---|---|---|---|---|---|
| EPU T1 | 0.462*** | 0.273*** | -0.238** | -0.273** | -0.977*** | 1.439*** |
| | (2.79) | (3.24) | (-2.51) | (-2.07) | (-4.10) | (4.22) |
| EPU T2 | 0.328*** | 0.077 | -0.125 | -0.276* | -0.484** | 0.812*** |
| | (3.44) | (0.54) | (-1.47) | (-1.83) | (-2.45) | (3.15) |
| EPU T3 | 0.151 | 0.140* | 0.125 | 0.042 | 0.010 | 0.142 |
| | (1.02) | (1.71) | (0.95) | (0.23) | (0.04) | (0.42) |
| EPU T1-T3 | 0.310 | 0.133 | -0.363** | -0.315 | -0.986*** | 1.297*** |
| | (1.40) | (1.15) | (-2.28) | (-1.41) | (-2.90) | (2.73) |

Panel B: Mega Cap

| | EHB Q1 | EHB Q2 | EHB Q3 | EHB Q4 | EHB Q5 | EHB Q1-Q5 |
|---|---|---|---|---|---|---|
| EPU T1 | 0.573*** | 0.132 | -0.019 | -0.178** | -0.157 | 0.730*** |
| | (3.39) | (1.05) | (-0.19) | (-1.96) | (-0.97) | (2.62) |
| EPU T2 | 0.123 | 0.181* | -0.127 | 0.154 | -0.386** | 0.510* |
| | (0.79) | (1.92) | (-1.10) | (1.36) | (-2.10) | (1.73) |
| EPU T3 | 0.082 | -0.137 | 0.001 | 0.165 | -0.028 | 0.110 |
| | (0.63) | (-1.14) | (0.01) | (1.43) | (-0.14) | (0.36) |
| EPU T1-T3 | 0.491** | 0.269 | -0.020 | -0.343** | -0.129 | 0.620 |
| | (2.30) | (1.55) | (-0.13) | (-2.35) | (-0.50) | (1.50) |

Table 3.7: IVOL$_{MA36}$ and the Return Predictability of Analysts' Conditional Biases

This table presents the Fama French Five Factor alphas for double sort portfolios created by cross-sectionally sorting companies into terciles based on IVOL$_{MA36}$ then conditionally sorting firms into quintiles based on their EHB. Panel A shows the results for non-mega-cap firms and panel B shows the results for mega-cap firms. Firms are split into mega- and non-mega-cap firms prior to creating the portfolios. Portfolios are value weighted and are re-balanced on a monthly basis. Standard errors of the resulting regression coefficients are computed based on Newey and West, 1987 with 12 lags. Statistical significance is denoted as ***, **, and * for $p<0.10$, $p<0.05$, and $p<0.01$, respectively. T1 indicates the lowest values and T3 the highest values for IVOL$_{MA36}$. Q1 indicates the lowest values and Q5 the highest values for EHB. The sample period is from January 1996 to December 2019.

Panel A: Non-Mega Cap

| | EHB Q1 | EHB Q2 | EHB Q3 | EHB Q4 | EHB Q5 | EHB Q1-Q5 |
|---|---|---|---|---|---|---|
| IVOL$_{MA36}$ T1 | -0.062 | 0.090 | -0.003 | -0.033 | 0.060 | -0.123 |
| | (-0.57) | (0.97) | (-0.03) | (-0.44) | (0.40) | (-0.66) |
| IVOL$_{MA36}$ T2 | 0.042 | 0.097 | -0.233** | -0.257** | -0.463** | 0.505* |
| | (0.39) | (1.16) | (-2.20) | (-2.44) | (-2.10) | (1.90) |
| IVOL$_{MA36}$ T3 | 0.457** | 0.350** | -0.027 | -0.336 | -0.540** | 0.998*** |
| | (2.22) | (2.04) | (-0.18) | (-1.59) | (-2.10) | (2.58) |
| IVOL$_{MA36}$ T1-T3 | -0.520** | -0.260 | 0.024 | 0.303 | 0.601*** | -1.120*** |
| | (-2.00) | (-1.16) | (0.12) | (1.44) | (2.62) | (-3.06) |

Panel B: Mega Cap

| | EHB Q1 | EHB Q2 | EHB Q3 | EHB Q4 | EHB Q5 | EHB Q1-Q5 |
|---|---|---|---|---|---|---|
| IVOL$_{MA36}$ T1 | -0.131 | -0.057 | -0.126 | -0.031 | -0.004 | -0.128 |
| | (-1.10) | (-0.61) | (-1.11) | (-0.23) | (-0.03) | (-0.66) |
| IVOL$_{MA36}$ T2 | 0.093 | 0.029 | -0.113 | 0.072 | -0.051 | 0.144 |
| | (0.72) | (0.25) | (-0.98) | (0.62) | (-0.31) | (0.60) |
| IVOL$_{MA36}$ T3 | 0.458** | 0.388** | 0.135 | 0.071 | 0.012 | 0.447 |
| | (1.97) | (2.24) | (1.14) | (0.40) | (0.05) | (1.38) |
| IVOL$_{MA36}$ T1-T3 | -0.590* | -0.445** | -0.262 | -0.102 | -0.016 | -0.574 |
| | (-1.93) | (-2.30) | (-1.51) | (-0.51) | (-0.05) | (-1.35) |

Table 3.8: Abnormal IVOL and the Return Predictability of Analysts' Conditional Biases

This table presents the Fama French Five Factor alphas for double sort portfolios created by cross-sectionally sorting companies into terciles based on Abnormal IVOL then conditionally cross-sectionally sorting firms into quintiles based on their EHB. Firms are split into mega- and non-mega-cap firms prior to creating the portfolios. Panel A shows the results for non-mega-cap firms and panel B shows the results for mega-cap firms. Portfolios are value weighted and are re-balanced on a monthly basis. Standard errors of the resulting regression coefficients are computed based on Newey and West, 1987 with 12 lags. Statistical significance is denoted as ***, **, and * for $p<0.10$, $p<0.05$, and $p<0.01$, respectively. T1 indicates the lowest values and T3 the highest values for Abnormal IVOL. Q1 indicates the lowest values and Q5 the highest values for EHB. The sample period is from June 1990 to December 2019.

Panel A: Non-Mega Cap

| | EHB Q1 | EHB Q2 | EHB Q3 | EHB Q4 | EHB Q5 | EHB Q1-Q5 |
|---|---|---|---|---|---|---|
| Abnormal IVOL T1 | 0.347*** | 0.277*** | -0.104 | -0.087 | -0.363* | 0.710*** |
| | (3.26) | (3.00) | (-1.09) | (-0.87) | (-1.92) | (2.83) |
| Abnormal IVOL T2 | 0.100 | 0.064 | 0.127 | -0.044 | -0.289 | 0.389 |
| | (0.85) | (0.80) | (1.41) | (-0.37) | (-1.53) | (1.52) |
| Abnormal IVOL T3 | 0.131 | -0.031 | -0.184* | -0.097 | -0.368 | 0.499 |
| | (1.04) | (-0.35) | (-1.66) | (-0.67) | (-1.45) | (1.56) |
| Abnormal IVOL T1-T3 | 0.217* | 0.308** | 0.079 | 0.010 | 0.005 | 0.212 |
| | (1.78) | (2.05) | (0.51) | (0.06) | (0.02) | (0.88) |

Panel B: Mega Cap

| | EHB Q1 | EHB Q2 | EHB Q3 | EHB Q4 | EHB Q5 | EHB Q1-Q5 |
|---|---|---|---|---|---|---|
| Abnormal IVOL T1 | 0.504*** | 0.360** | 0.149 | 0.168 | 0.027 | 0.476* |
| | (2.82) | (2.46) | (1.11) | (1.21) | (0.18) | (1.94) |
| Abnormal IVOL T2 | -0.025 | -0.047 | -0.104 | 0.033 | -0.003 | -0.022 |
| | (-0.18) | (-0.38) | (-0.87) | (0.28) | (-0.02) | (-0.10) |
| Abnormal IVOL T3 | -0.280* | -0.065 | -0.137 | -0.031 | -0.230 | -0.050 |
| | (-1.84) | (-0.48) | (-0.92) | (-0.19) | (-1.02) | (-0.19) |
| Abnormal IVOL T1-T3 | 0.783*** | 0.425** | 0.286 | 0.199 | 0.257 | 0.526 |
| | (2.65) | (2.10) | (1.27) | (0.83) | (0.89) | (1.45) |

Table 3.9: Firm Age and the Return Predictability of Analysts' Conditional Biases

This table presents the Fama French Five Factor alphas for double sort portfolios created by cross-sectionally sorting companies into terciles based on 1/Age then conditionally cross-sectionally sorting firms into quintiles based on their EHB. Firms are split into mega- and non-mega-cap firms prior to creating the portfolios. Panel A shows the results for non-mega-cap firms and panel B shows the results for mega-cap firms. Portfolios are value weighted and are re-balanced on a monthly basis. Standard errors of the resulting regression coefficients are computed based on Newey and West, 1987 with 12 lags. Statistical significance is denoted as ***, **, and * for $p<0.10$, $p<0.05$, and $p<0.01$, respectively. T1 indicates the highest values and T3 the lowest values for Firm Age. Q1 indicates the lowest values and Q5 the highest values for EHB. The sample period is from June 1990 to December 2019.

Panel A: Non-Mega Cap

| | EHB Q1 | EHB Q2 | EHB Q3 | EHB Q4 | EHB Q5 | EHB Q1-Q5 |
|---|---|---|---|---|---|---|
| 1/Age T1 | -0.087 | 0.020 | -0.006 | -0.113 | -0.191 | 0.104 |
| | (-0.81) | (0.31) | (-0.09) | (-1.09) | (-1.00) | (0.41) |
| 1/Age T2 | 0.315** | 0.231*** | -0.059 | 0.006 | -0.349 | 0.664** |
| | (2.53) | (3.26) | (-0.66) | (0.05) | (-1.64) | (2.25) |
| 1/Age T3 | 0.580*** | 0.291*** | -0.135 | -0.113 | -0.694*** | 1.274*** |
| | (3.63) | (3.31) | (-1.32) | (-0.76) | (-3.26) | (4.03) |
| 1/Age T1-T3 | -0.667*** | -0.271*** | 0.128 | 0.000 | 0.504*** | -1.170*** |
| | (-3.72) | (-2.60) | (1.07) | (0.00) | (2.76) | (-4.43) |

Panel B: Mega Cap

| | EHB Q1 | EHB Q2 | EHB Q3 | EHB Q4 | EHB Q5 | EHB Q1-Q5 |
|---|---|---|---|---|---|---|
| 1/Age T1 | -0.027 | -0.104 | -0.064 | 0.120 | -0.202 | 0.176 |
| | (-0.24) | (-0.98) | (-0.79) | (1.20) | (-1.39) | (0.87) |
| 1/Age T2 | 0.018 | 0.077 | -0.083 | -0.018 | 0.046 | -0.027 |
| | (0.15) | (0.67) | (-0.68) | (-0.16) | (0.22) | (-0.10) |
| 1/Age T3 | 0.513** | 0.268* | 0.328** | 0.087 | -0.142 | 0.655* |
| | (2.19) | (1.95) | (2.24) | (0.72) | (-0.82) | (1.92) |
| 1/Age T1-T3 | -0.540** | -0.371** | -0.392** | 0.033 | -0.061 | -0.479 |
| | (-2.14) | (-2.21) | (-2.35) | (0.20) | (-0.36) | (-1.60) |

Table 3.10: Information Cost and the Return Predictability of Analysts' Conditional Biases: Full Sample

This table presents the results of Fama-MacBeth monthly regressions of one-month-ahead returns (in percentage points) on the normalized rank of EHB (i.e., the rank scaled by the number of stocks in the cross-section) by 1/Age terciles. The 1/Age terciles are generated by cross-sectionally sorting firms into terciles based on 1/Age. Columns 2, 4, 6, and 8 include Operating Profitability (Revenue minus cost - administrative expenses - interest expenses, scaled by book value of equity), Asset Growth (proxy for Investments), BTM, 6 month Momentum, and Size as controls. Controls are also the normalized rank (i.e., the rank scaled by the number of stocks in the cross-section) by tercile. Standard errors of the resulting Fama-MacBeth regression coefficient are computed based on Newey and West, 1987 with 12 lags. Statistical significance is denoted as ***, **, and * for $p<0.10$, $p<0.05$, and $p<0.01$, respectively. T1 indicates the oldest firms and T3 the youngest firms. The sample period is from June 1990 to December 2019.

| | T1 | | T2 | | T3 | | T3-T1 | |
|---|---|---|---|---|---|---|---|---|
| EHB | 0.003 | -0.002 | -0.004* | -0.006** | -0.014*** | -0.009*** | -0.017*** | -0.007*** |
| | (1.2) | (-0.9) | (-1.7) | (-2.1) | (-4.3) | (-3.4) | (-5.5) | (-3.3) |
| Op. Profit. | | 0.002 | | 0.004** | | 0.006** | | 0.004* |
| | | (1.1) | | (2.0) | | (2.3) | | (1.8) |
| Asset Growth | | 0.000 | | 0.005*** | | 0.008*** | | 0.008*** |
| | | (0.2) | | (2.9) | | (3.7) | | (3.4) |
| BTM | | 0.003 | | 0.004 | | 0.002 | | -0.002 |
| | | (1.1) | | (0.8) | | (0.3) | | (-0.5) |
| 6m Momentum | | -0.004 | | 0.002 | | 0.009** | | 0.012*** |
| | | (-1.6) | | (0.9) | | (2.5) | | (4.4) |
| Size | | -0.002 | | -0.003* | | -0.003 | | -0.001 |
| | | (-1.1) | | (-1.8) | | (-1.3) | | (-0.6) |
| Observations | 216708 | 213045 | 208591 | 201969 | 209929 | 147627 | 426637 | 360672 |
| Adjusted $R^2$ | 0.020 | 0.066 | 0.013 | 0.051 | 0.012 | 0.050 | 0.030 | 0.069 |

Table 3.11: Information Cost and the Return Predictability of Analysts' Conditional Biases

This table presents the results of Fama-MacBeth monthly regressions of one-month-ahead returns (in percentage points) on the normalized rank of EHB (i.e., the rank scaled by the number of stocks in the cross-section) by 1/Age terciles for earnings announcement and non-earnings announcement months. The 1/Age terciles are generated by cross-sectionally sorting firms into terciles based on 1/Age. Columns 2, 4, 6, and 8 include Operating Profitability (Revenue minus cost - administrative expenses - interest expenses, scaled by book value of equity), Asset Growth (proxy for Investments), BTM, 6 month Momentum, and Size as controls. Controls are also the normalized rank (i.e., the rank scaled by the number of stocks in the cross-section) by tercile. Panel A includes firm-months that do not have earnings announcements within the month. Panel B includes only firm-months that have earnings announcements within the month. Panel C provides the difference between months with and without earnings announcements. Standard errors of the resulting Fama-MacBeth regression coefficient are computed based on Newey and West, 1987 with 12 lags. Statistical significance is denoted as ***, **, and * for $p<0.10$, $p<0.05$, and $p<0.01$, respectively. T1 indicates the oldest firms and T3 the youngest firms. The sample period is from June 1990 to December 2019.

Panel A: Non-Earnings Announcement Months

| | T1 | | T2 | | T3 | | T3-T1 | |
|---|---|---|---|---|---|---|---|---|
| EHB | 0.003 | -0.003 | -0.008*** | -0.010*** | -0.019*** | -0.014*** | -0.022*** | -0.011*** |
| | (0.9) | (-1.0) | (-3.2) | (-3.4) | (-5.0) | (-4.8) | (-5.5) | (-4.1) |
| Op. Profit. | | 0.002 | | 0.005** | | 0.007** | | 0.005 |
| | | (0.9) | | (2.1) | | (2.2) | | (1.6) |
| Asset Growth | | 0.001 | | 0.005** | | 0.009*** | | 0.008*** |
| | | (0.9) | | (2.5) | | (3.8) | | (3.0) |
| BTM | | 0.004 | | 0.004 | | 0.003 | | -0.001 |
| | | (1.3) | | (1.0) | | (0.6) | | (-0.4) |
| 6m Momentum | | -0.004 | | 0.002 | | 0.011*** | | 0.015*** |
| | | (-1.5) | | (0.6) | | (2.6) | | (4.6) |
| Size | | 0.000 | | -0.003 | | -0.002 | | -0.002 |
| | | (0.1) | | (-1.6) | | (-0.9) | | (-1.0) |
| Observations | 145591 | 143121 | 140206 | 135671 | 140981 | 99440 | 286572 | 242561 |
| Adjusted $R^2$ | 0.020 | 0.066 | 0.014 | 0.050 | 0.013 | 0.050 | 0.030 | 0.069 |

Panel B: Earnings Announcement Months

| | T1 | | T2 | | T3 | | T3-T1 | |
|---|---|---|---|---|---|---|---|---|
| EHB | 0.003 | -0.003 | -0.004 | -0.003 | -0.008** | -0.001 | -0.011*** | 0.002 |
| | (1.1) | (-1.3) | (-1.1) | (-1.0) | (-2.5) | (-0.3) | (-3.2) | (0.4) |
| Op. Profit. | | 0.003 | | -0.002 | | 0.007 | | 0.003 |
| | | (1.4) | | (-0.6) | | (1.3) | | (0.7) |
| Asset Growth | | -0.001 | | 0.005* | | 0.010** | | 0.011** |
| | | (-0.6) | | (1.7) | | (2.2) | | (2.3) |
| BTM | | 0.006 | | 0.004 | | 0.000 | | -0.005 |
| | | (1.6) | | (0.8) | | (0.1) | | (-0.8) |
| 6m Momentum | | -0.001 | | 0.002 | | 0.006 | | 0.007 |
| | | (-0.4) | | (0.9) | | (1.3) | | (1.4) |
| Size | | -0.004 | | 0.001 | | 0.003 | | 0.006 |
| | | (-1.3) | | (0.4) | | (0.6) | | (1.2) |
| Observations | 71117 | 69924 | 68385 | 66298 | 68881 | 48145 | 139998 | 118069 |
| Adjusted $R^2$ | 0.019 | 0.080 | 0.020 | 0.065 | 0.013 | 0.054 | 0.032 | 0.091 |

Panel C: Difference Between Earnings Announcement and Non-Earnings Announcement Months

| | T1 | | T2 | | T3 | | T3-T1 | |
|---|---|---|---|---|---|---|---|---|
| EHB | -0.000 | 0.001 | -0.005** | -0.007** | -0.011*** | -0.012*** | -0.010** | -0.013** |
| | (-0.1) | (0.2) | (-2.0) | (-2.4) | (-3.1) | (-3.0) | (-2.5) | (-2.5) |
| Op. Profit. | | -0.002 | | 0.007** | | -0.000 | | 0.001 |
| | | (-0.7) | | (2.5) | | (-0.1) | | (0.2) |
| Asset Growth | | 0.002 | | 0.001 | | -0.000 | | -0.003 |
| | | (1.3) | | (0.2) | | (-0.1) | | (-0.6) |
| BTM | | -0.001 | | 0.001 | | 0.003 | | 0.004 |
| | | (-0.5) | | (0.2) | | (0.5) | | (0.6) |
| 6m Momentum | | -0.003 | | -0.001 | | 0.005 | | 0.008 |
| | | (-1.2) | | (-0.2) | | (0.8) | | (1.3) |
| Size | | 0.004 | | -0.004 | | -0.005 | | -0.009 |
| | | (1.6) | | (-1.4) | | (-1.2) | | (-1.6) |
| Observations | 216708 | 213045 | 208591 | 201969 | 209862 | 147585 | 426570 | 360630 |
| Adjusted $R^2$ | 0.023 | 0.072 | 0.016 | 0.055 | 0.013 | 0.053 | 0.032 | 0.074 |

Table 3.12: Information Demand

This table runs a probit regression to analyze the relationship between AIA and firm uncertainty. We use a probit model to regress the daily AIA on OIV, Abnormal OIV, HLtH, or Abnormal HLtH. OIV and Abnormal OIV is winsorized at the 1 and 99th percentile. HLtH and Abnormal HLtH is winsorized at the 99th percentile. For days when Bloomberg has no value, we do not fill the missing information of AIA with zero. Each regression includes a control for Size. Standard errors are clustered by firm and day. Statistical significance is denoted as ***, **, and * for $p<0.10$, $p<0.05$, and $p<0.01$, respectively. The sample period is from March 2010 to December 2019.

| | (1) | (2) | (3) | (4) |
|---|---|---|---|---|
| OIV | 1.209*** | | | |
| | (30.4) | | | |
| Ab. OIV | | 0.550*** | | |
| | | (21.9) | | |
| HLtH | | | 25.620*** | |
| | | | (53.3) | |
| Ab. HLtH | | | | 0.921*** |
| | | | | (58.9) |
| Size | 0.235*** | 0.172*** | 0.284*** | 0.193*** |
| | (46.5) | (33.3) | (60.6) | (34.6) |
| Observations | 2647050 | 2646269 | 2681426 | 2681426 |
| Pseudo $R^2$ | 0.042 | 0.034 | 0.111 | 0.135 |

Table 3.13: Information Demand by Terciles

This table runs a probit regression to analyze the relationship between AIA and firm uncertainty. We use a probit model to regress the daily AIA on Abnormal OIV or Abnormal HLtH for firms sorted into long-run uncertainty terciles. The terciles are generated by orthogonalizing the cross-sectional normalized rank of $\text{IVOL}_{MA36}$ by to the cross-sectional normalized rank of Size. This is done on a monthly basis and firms are sorted into terciles based on the values generated at the end of the previous month. Abnormal OIV is winsorized at the 1 and 99th percentile. Abnormal HLtH is winsorized at the 99th percentile. For days when Bloomberg has no value, we do not fill the missing information of AIA with zero. Each regression includes a control for Size. Panel A uses the raw values of the uncertainty measures and Size. Panel B used the cross-sectional normalized rank (i.e., the rank scaled by the number of stocks in the cross-section) of the uncertainty measures and Size. Standard errors are clustered by firm and day. Statistical significance is denoted as ***, **, and * for p<0.10, p<0.05, and p<0.01, respectively. The sample period is from March 2010 to December 2019.

Panel A: Raw Values

| | (1) T1 | (2) T2 | (3) T3 | (4) T1-T3 | (5) T1 | (6) T2 | (7) T3 | (8) T1-T3 |
|---|---|---|---|---|---|---|---|---|
| Ab. OIV | 0.426*** | 0.577*** | 0.630*** | -0.204*** | | | | |
| | (14.0) | (18.6) | (20.9) | (-6.3) | | | | |
| Ab. HLtH | | | | | 0.789*** | 0.919*** | 1.043*** | -0.254*** |
| | | | | | (46.8) | (56.9) | (61.9) | (-19.1) |
| Size | 0.200*** | 0.193*** | 0.175*** | 0.025*** | 0.216*** | 0.216*** | 0.200*** | 0.016 |
| | (28.5) | (29.8) | (23.4) | (2.6) | (29.0) | (30.8) | (24.2) | (1.5) |
| Observations | 803138 | 878222 | 964909 | 1768047 | 819325 | 889614 | 972487 | 1791812 |
| Pseudo $R^2$ | 0.049 | 0.042 | 0.028 | 0.039 | 0.124 | 0.145 | 0.150 | 0.141 |

Panel B: Normalized Rank

| | (1) T1 | (2) T2 | (3) T3 | (4) T1-T3 | (5) T1 | (6) T2 | (7) T3 | (8) T1-T3 |
|---|---|---|---|---|---|---|---|---|
| Ab. OIV Normalized Rank | 0.242*** | 0.294*** | 0.355*** | -0.113*** | | | | |
| | (16.8) | (20.9) | (28.1) | (-6.5) | | | | |
| Ab. HLtH Normalized Rank | | | | | 1.137*** | 1.397*** | 1.611*** | -0.475*** |
| | | | | | (57.6) | (75.6) | (86.5) | (-19.8) |
| Size Normalized Rank | 1.087*** | 0.952*** | 0.771*** | 0.316*** | 1.146*** | 1.025*** | 0.844*** | 0.302*** |
| | (25.3) | (27.0) | (23.1) | (6.0) | (25.8) | (27.3) | (23.2) | (5.4) |
| Observations | 803138 | 878222 | 964909 | 1768047 | 819325 | 889614 | 972487 | 1791812 |
| Pseudo $R^2$ | 0.050 | 0.041 | 0.030 | 0.040 | 0.099 | 0.111 | 0.119 | 0.112 |

Table 3.14: EHB and Analysts' Revision Correlations by Idiosyncratic Volatility Tercile

This table shows the Pearson correlation between EHB and analysts' revisions across the terciles of IVOL. EHB is the average rank of the EHB across the four forecasting horizons.

| | Spearman Correlation |
|---|---|
| IVOL T1 | -0.599 |
| IVOL T2 | -0.612 |
| IVOL T3 | -0.567 |

Table 3.15: Return Predictability: Analysts' Revision

This table presents the results of Fama-MacBeth monthly regressions of one-month-ahead returns (in percentage points) on the normalized rank of EHB (i.e., the rank scaled by the number of stocks in the cross-section) and a Revision dummy by IVOL tercile.The revision dummy is equal to 1 when the average 3-month revision across the four horizons is greater than zero, 0 when the revision is equal to zero, and -1 otherwise. The IVOL terciles are generated by cross-sectionally sorting firms into terciles based on IVOL. Standard errors are computed based on Newey and West, 1987 with 12 lags. Statistical significance is denoted as ***, **, and * for $p<0.10$, $p<0.05$, and $p<0.01$, respectively.The sample period is from June 1990 to December 2019.

| | T1 | | T3 | | T1-T3 | |
|---|---|---|---|---|---|---|
| Revision Dummy | 0.007 | 0.028 | 0.165** | 0.063 | 0.159** | 0.035 |
| | (0.2) | (1.0) | (2.4) | (0.9) | (2.5) | (0.5) |
| EHB | | 0.002 | | -0.012*** | | -0.014*** |
| | | (1.4) | | (-3.1) | | (-3.9) |
| Observations | 223091 | 208619 | 222796 | 209912 | 445887 | 418531 |
| Adjusted $R^2$ | 0.006 | 0.018 | 0.004 | 0.013 | 0.042 | 0.051 |

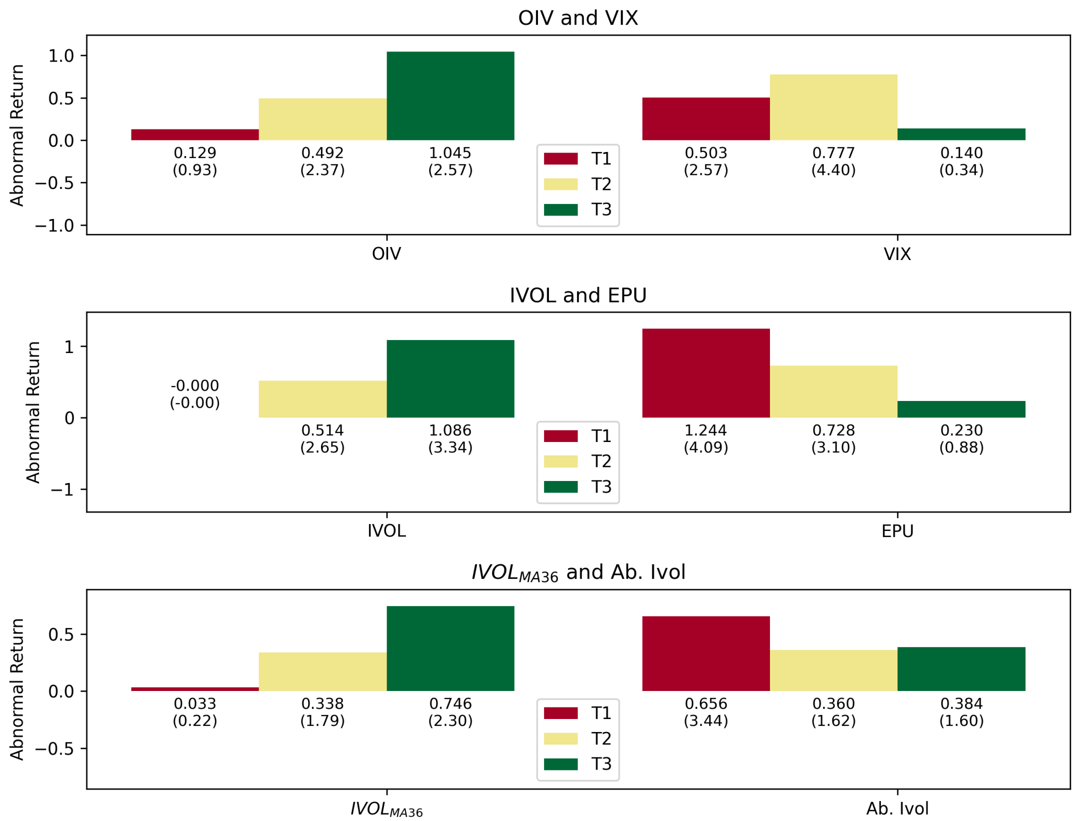



Figure 3.1: Return Predictability by Uncertainty Terciles

This figure shows the Fama French Five Factor alphas of the EHB Q1-Q5 portfolios by uncertainty terciles using our six monthly uncertainty measures. The EHB Q1-Q5 portfolios based on OIV , IVOL, IVOL$_{MA36}$, and abnormal IVOL, are made by cross-sectionally sorting companies first into deciles based on their market value of equity, then conditionally cross-sectionally sorting firms into terciles based on each uncertainty measure. Then firms are then conditionally cross-sectionally sorted into quintiles based on EHB. The EHB Q1-Q5 portfolios based on VIX, and EPU are made by sorting observations into terciles in the time series based on each uncertainty measure. Then, within each cross section, we first divide firms into the 3 size groups from Table 3.2. We finally conditionally cross-sectionally sort firms into quintiles based on EHB. Within each cross section, we calculate the value weighted return within each size-EHB group, and then average across the three size groups. Portfolios are value weighted and are re-balanced on a monthly basis. Q1 (Q5) contains firms with the lowest (highest) values of EHB. T1 (T3) of each uncertainty tercile contains firms with the lowest (highest) values. The average value and test statistics are in presented below their corresponding bars. Standard errors of the resulting regression coefficients are computed based on Newey and West, 1987 with 12 lags. Statistical significance is denoted as ***, **, and * for $p<0.10$, $p<0.05$, and $p<0.01$, respectively. The sample period is from June 1990 to December 2019, figures. with the exception of OIV which begins in January 1996. Please see

# Bibliography


Acharya, V. V., Berger, A. N., & Roman, R. A. (2018). Lending implications of U.S. bank stress tests: Costs or benefits? *Journal of Financial Intermediation*, *34*, 58–90. https://doi.org/10.1016/j.jfi.2018.01.004

Ali, A., Hwang, L.-S., & Trombley, M. A. (2003). Residual Income Based Valuation Predicts Future Stock Returns: Evidence on Mispricing vs. Risk Explanations. *The Accounting Review*, *78*(2), 377–396. https://doi.org/10.2308/accr.2003.78.2.377

Altamuro, J., & Beatty, A. (2010). How does internal control regulation affect financial reporting? *Journal of Accounting and Economics*, *49*(1-2), 58–74. https://doi.org/10.1016/j.jacceco.2009.07.002

Alvero, A., Ando, S., & Xiao, K. (2022). Watch What They Do, Not What They Say: Estimating Regulatory Costs from Revealed Preferences. *IMF Working Papers*, *22*(41).

Andrei, D., Friedman, H. L., & Ozel, N. B. (2021). Economic Uncertainty and Investor Attention. https://doi.org/10.2139/ssrn.3128673

Andreou, P. C., Philip, D., & Robejsek, P. (2016). Bank Liquidity Creation and Risk-Taking: Does Managerial Ability Matter?: BANK LIQUIDITY CREATION AND

RISK-TAKING. *Journal of Business Finance & Accounting*, *43*(1-2), 226–259. https://doi.org/10.1111/jbfa.12169

Baker, S. R., Bloom, N., & Davis, S. J. (2016). Measuring Economic Policy Uncertainty. *The Quarterly Journal of Economics*, *131*(4), 1593–1636. https://doi.org/10.1093/qje/qjw024

Baley, I., & Veldkamp, L. (2021). Bayesian Learning. Retrieved September 15, 2022, from https://papers.ssrn.com/abstract=3928655

Ballew, H., Iselin, M., & Nicoletti, A. (2022). Accounting-based thresholds and growth decisions in the banking industry. *Review of Accounting Studies*, *27*(1), 232–274. https://doi.org/10.1007/s11142-021-09589-6

Benamar, H., Foucault, T., & Vega, C. (2021). Demand for Information, Uncertainty, and the Response of U.S. Treasury Securities to News. *The Review of Financial Studies*, *34*(7), 3403–3455. https://doi.org/10.1093/rfs/hhaa072

Ben-Rephael, A., Da, Z., & Israelsen, R. D. (2017). It Depends on Where You Search: Institutional Investor Attention and Underreaction to News. *The Review of Financial Studies*, *30*(9), 3009–3047. https://doi.org/10.1093/rfs/hhx031

Berger, A. N., & Bouwman, C. H. S. (2009). Bank Liquidity Creation. *Review of Financial Studies*, *22*(9), 3779–3837. https://doi.org/10.1093/rfs/hhn104

Berger, A. N., Bouwman, C. H. S., & Kim, D. (2017). Small Bank Comparative Advantages in Alleviating Financial Constraints and Providing Liquidity Insurance over Time. *The Review of Financial Studies*, *30*(10), 3416–3454. https://doi.org/10.1093/rfs/hhx038

Berger, A. N., Bouwman, C. H., Kick, T., & Schaeck, K. (2016). Bank liquidity creation following regulatory interventions and capital support. *Journal of Financial Intermediation*, *26*, 115–141. https://doi.org/10.1016/j.jfi.2016.01.001

Berger, A. N., & Sedunov, J. (2017). Bank liquidity creation and real economic output. *Journal of Banking & Finance*, *81*, 1–19. https://doi.org/10.1016/j.jbankfin.2017.04.005

Bhattacharya, S., & Thakor, A. V. (1993). Contemporary Banking Theory. *Journal of Financial Intermediation*, *3*(1), 2–50. https://doi.org/10.1006/jfin.1993.1001

Bindal, S., Bouwman, C. H., Hu, S. (, & Johnson, S. A. (2020). Bank regulatory size thresholds, merger and acquisition behavior, and small business lending. *Journal of Corporate Finance*, *62*, 101519. https://doi.org/10.1016/j.jcorpfin.2019.101519

Blankespoor, E., deHaan, E., & Marinovic, I. (2020). Disclosure processing costs, investors information choice, and equity market outcomes: A review. *Journal of Accounting and Economics*, *70*(2-3), 101344. https://doi.org/10.1016/j.jacceco.2020.101344

Bostic, R., & Lee, H. (2017). Small Business Lending Under the Community Reinvestment Act. *Cityscape*, *19*(2).

Bouwman, C. H., Hu, S. (, & Johnson, S. A. (2018). Differential bank behaviors around the Dodd–Frank Act size thresholds. *Journal of Financial Intermediation*, *34*, 47–57. https://doi.org/10.1016/j.jfi.2018.01.005

Breiman, L. (1996). Bagging predictors. *Machine Learning*, *24*(2), 123–140. https://doi.org/10.1007/BF00058655

Breitenstein, E., Cofer, R., DiChiara, R., French, G., Hanrahan, M., Hoople, D., Polson, C., & Raslavich, C. (2020). *FDIC Community Banking Study* (tech. rep.). Federal Deposit Insurance Corporation.

Bryant, J. (1980). A model of reserves, bank runs, and deposit insurance. *Journal of Banking & Finance*, *4*(4), 335–344. https://doi.org/10.1016/0378-4266(80)90012-6

Casu, B., di Pietro, F., & Trujillo-Ponce, A. (2019). Liquidity Creation and Bank Capital. *Journal of Financial Services Research*, *56*(3), 307–340. https://doi.org/10.1007/s10693-018-0304-y

Chatterjee, U. K. (2018). Bank liquidity creation and recessions. *Journal of Banking & Finance*, *90*, 64–75. https://doi.org/10.1016/j.jbankfin.2018.03.002

Chen, A. Y., & Zimmermann, T. (2022). Open Source Cross-Sectional Asset Pricing. *Critical Finance Review*, 69.

Clark, E., Degl'Innocenti, M., Radic, N., & Zhou, S. (2020). Heterogeneous Risk Behaviour of Bank Holding Companies after the Dodd–Frank Act. *SSRN Electronic Journal*. https://doi.org/10.2139/ssrn.3625670

Cyree, K. B. (2016). The Effects of Regulatory Compliance for Small Banks Around Crisis-Based Regulation:Effects of Regulatory Compliance. *Journal of Financial Research*, *39*(3), 215–246. https://doi.org/10.1111/jfir.12096

Da, Z., Engelberg, J., & Gao, P. (2011). In Search of Attention. *The Journal of Finance*, *66*(5), 1461–1499. https://doi.org/10.1111/j.1540-6261.2011.01679.x

Dahl, D., & Franke, M. (2017). "Banking Deserts" Become a Concern as Branches Dry Up. *Regional Economist*. https://www.stlouisfed.org/publications/regional-economist/second-quarter-2017/banking-deserts-become-a-concern-as-branches-dry-up

Dang, V. D., & Dang, V. C. (2021). How do bank characteristics affect the bank liquidity creation channel of monetary policy? *Finance Research Letters*, *43*, 101984. https://doi.org/10.1016/j.frl.2021.101984

Dávila, E., & Parlatore, C. (2023). Volatility and informativeness. *Journal of financial economics*, *147*(3), 550–572.

Davydov, D., Vähämaa, S., & Yasar, S. (2021). Bank liquidity creation and systemic risk. *Journal of Banking & Finance*, *123*, 106031. https://doi.org/10.1016/j.jbankfin.2020.106031

Dellavigna, S., & Pollet, J. M. (2009). Investor Inattention and Friday Earnings Announcements [_eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1111/j.1540-6261.2009.01447.x]. *The Journal of Finance*, *64*(2), 709–749. https://doi.org/10.1111/j.1540-6261.2009.01447.x

de Silva, T., & Thesmar, D. (2021). *Noise in Expectations: Evidence from Analyst Forecasts* (tech. rep. w28963). National Bureau of Economic Research. Cambridge, MA. https://doi.org/10.3386/w28963

Diamond, D. W., & Dybvig, P. H. (1983). Bank Runs, Deposit Insurance, and Liquidity. *Journal of Political Economy*, *91*(3), 401–419. https://doi.org/10.1086/261155

Díaz, V., & Huang, Y. (2017). The role of governance on bank liquidity creation. *Journal of Banking & Finance*, *77*, 137–156. https://doi.org/10.1016/j.jbankfin.2017.01.003

Diether, K. B., Malloy, C. J., & Scherbina, A. (2002). Differences of Opinion and the Cross Section of Stock Returns. *The Journal of Finance*, *57*(5), 2113–2141. https://doi.org/10.1111/0022-1082.00490

Ding, L., Lee, H., & Bostic, R. W. (2020). Effects of the Community Reinvestment Act on small business lending. *Journal of Urban Affairs*, 1–20. https://doi.org/10.1080/07352166.2020.1808005

Duan, Y., & Niu, J. (2020). Liquidity creation and bank profitability. *The North American Journal of Economics and Finance*, *54*, 101250. https://doi.org/10.1016/j.najef.2020.101250

Fama, E. F. (1970). Efficient Capital Markets: A Review of Theory and Empirical Work. *The Journal of Finance*, *25*(2), 383. https://doi.org/10.2307/2325486

Faulkender, M., & Petersen, M. (2012). Investment and Capital Constraints: Repatriations Under the American Jobs Creation Act. *Review of Financial Studies*, *25*(11), 3351–3388. https://doi.org/10.1093/rfs/hhs092

Feng, G., Giglio, S., & Xiu, D. (2020). Taming the Factor Zoo: A Test of New Factors. *The Journal of Finance*, *75*(3), 1327–1370. https://doi.org/10.1111/jofi.12883

Frankel, R., & Lee, C. M. (1998). Accounting valuation, market expectation, and cross-sectional stock returns. *Journal of Accounting and Economics*, *25*(3), 283–319. https://doi.org/10.1016/S0165-4101(98)00026-3

Frazier, P. I. (2018). A Tutorial on Bayesian Optimization [arXiv: 1807.02811]. *arXiv:1807.02811 [cs, math, stat]*. Retrieved January 21, 2022, from http://arxiv.org/abs/1807.02811

Friedman, J. H. (2001). Greedy function approximation: A gradient boosting machine. *The Annals of Statistics*, *29*(5). https://doi.org/10.1214/aos/1013203451

Friedman, J. H. (2002). Stochastic gradient boosting. *Computational Statistics & Data Analysis*, *38*(4), 367–378. https://doi.org/10.1016/S0167-9473(01)00065-2

Fungáčová, Z., Weill, L., & Zhou, M. (2017). Bank Capital, Liquidity Creation and Deposit Insurance. *Journal of Financial Services Research*, *51*(1), 97–123. https://doi.org/10.1007/s10693-016-0240-7

García, R. E., & Steele, S. (2022). Stress testing and bank business patterns: A regression discontinuity study. *Journal of Banking & Finance*, *135*, 105964. https://doi.org/10.1016/j.jbankfin.2020.105964

Gopalan, Y., Imdieke, A., Schroeder, J. H., & Stuber, S. (2020). Reliance on External Assurance in Bank Monitoring. *SSRN Electronic Journal*. https://doi.org/10.2139/ssrn.3440203

Green, J., Hand, J. R. M., & Zhang, X. F. (2017). The Characteristics that Provide Independent Information about Average U.S. Monthly Stock Returns. *The Review of Financial Studies*, *30*(12), 4389–4436. https://doi.org/10.1093/rfs/hhx019

Gu, S., Kelly, B., & Xiu, D. (2020). Empirical Asset Pricing via Machine Learning (A. Karolyi, Ed.). *The Review of Financial Studies*, *33*(5), 2223–2273. https://doi.org/10.1093/rfs/hhaa009

Hirshleifer, D. (2001). Investor Psychology and Asset Pricing. *The Journal of Finance*, *56*(4), 1533–1597. https://doi.org/10.1111/0022-1082.00379

Hirshleifer, D., Lim, S. S., & Teoh, S. H. (2009). Driven to Distraction: Extraneous Events and Underreaction to Earnings News. *The Journal of Finance*, *64*(5), 2289–2325. https://doi.org/10.1111/j.1540-6261.2009.01501.x

Holmström, B., & Tirole, J. (1998). Private and Public Supply of Liquidity. *Journal of Political Economy*, *106*(1), 1–40. https://doi.org/10.1086/250001

Horváth, R., Seidler, J., & Weill, L. (2014). Bank Capital and Liquidity Creation: Granger-Causality Evidence. *Journal of Financial Services Research*, *45*(3), 341–361. https://doi.org/10.1007/s10693-013-0164-4

Huang, Q. (2018). Impact of the Dodd-Frank Act on Systemic Risk: A Counterfactual Analysis. *SSRN Electronic Journal*. https://doi.org/10.2139/ssrn.3019977

Huang, S., Huang, Y., & Lin, T.-C. (2019). Attention allocation and return co-movement: Evidence from repeated natural experiments. *Journal of Financial Economics*, *132*(2), 369–383. https://doi.org/10.1016/j.jfineco.2018.10.006

Jiang, L. (2020). Risk Management Committee and Bank Performance: Evidence from the Adoption of Dodd-Frank Act. *SSRN Electronic Journal*. https://doi.org/10.2139/ssrn.3699957

Jiang, L., Levine, R., & Lin, C. (2019). Competition and Bank Liquidity Creation. *Journal of Financial and Quantitative Analysis*, *54*(2), 513–538. https://doi.org/10.1017/S0022109018000820

Jin, J. Y., Kanagaretnam, K., & Lobo, G. J. (2013). Unintended consequences of the increased asset threshold for FDICIA internal controls: Evidence from U.S. private banks. *Journal of Banking & Finance*, *37*(12), 4879–4892. https://doi.org/10.1016/j.jbankfin.2013.08.024

Jin, J. Y., Kanagaretnam, K., Lobo, G. J., & Mathieu, R. (2013). Impact of FDICIA internal controls on bank risk taking. *Journal of Banking & Finance*, *37*(2), 614–624. https://doi.org/10.1016/j.jbankfin.2012.09.013

Johnson, T. L., & So, E. C. (2018). Asymmetric Trading Costs Prior to Earnings Announcements: Implications for Price Discovery and Returns. *Journal of Accounting Research*, *56*(1), 217–263. https://doi.org/10.1111/1475-679X.12189

Kashyap, A. K., Rajan, R., & Stein, J. C. (2002). Banks as Liquidity Providers: An Explanation for the Coexistence of Lending and Deposit-taking. *The Journal of Finance*, *57*(1), 33–73. https://doi.org/10.1111/1540-6261.00415

Ke, G., Meng, Q., Finley, T., Wang, T., Chen, W., Ma, W., Ye, Q., & Liu, T.-Y. (2017). LightGBM: A Highly Efficient Gradient Boosting Decision Tree. *Advances in Neural Information Processing Systems*, *30*. Retrieved January 18, 2022, from https://proceedings.neurips.cc/paper/2017/hash/6449f44a102fde848669bdd9eb6b76fa-Abstract.html

Kim, M. J., Lee, K. M., & Earle, J. S. (2021). Does the Community Reinvestment Act increase lending to small businesses in lower income neighborhoods? *Economics Letters*, *209*, 110146. https://doi.org/10.1016/j.econlet.2021.110146

Kothari, S., So, E., & Verdi, R. (2016). Analysts' Forecasts and Asset Pricing: A Survey. *Annual Review of Financial Economics*, *8*(1), 197–219. https://doi.org/10.1146/annurev-financial-121415-032930

Kozak, S., Nagel, S., & Santosh, S. (2020). Shrinking the cross-section. *Journal of Financial Economics*, *135*(2), 271–292. https://doi.org/10.1016/j.jfineco.2019.06.008

Labonte, M., & Perkins, D. (2021). *Over the Line: Asset Thresholds in Bank Regulation* (CRS Report R46779). Congressional Research Service.

Liu, K. (2019). The Impact of the Dodd-Frank Act on Small U.S. Banks. *SSRN Electronic Journal*. https://doi.org/10.2139/ssrn.3419586

Loh, R. K., & Stulz, R. M. (2018). Is Sell-Side Research More Valuable in Bad Times? [_eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1111/jofi.12611]. *The Journal of Finance*, *73*(3), 959–1013. https://doi.org/10.1111/jofi.12611

Mackowiak, B., Matejka, F., & Wiederholt, M. (2021). Rational Inattention: A Review. *SSRN Electronic Journal*. https://doi.org/10.2139/ssrn.3871534

Modigliani, F., & Miller, M. H. (1958). The Cost of Capital, Corporation Finance and the Theory of Investment. *The American Economic Review*, *48*(3), 261–297. Retrieved March 3, 2014, from http://www.jstor.org/stable/1809766

Newey, W. K., & West, K. D. (1987). A simple, positive semi-definite, heteroskedasticity and autocorrelation consistent covariance matrix. *Econometrica*, *55*(3), 703. Retrieved September 17, 2015, from http://search.proquest.com/docview/214867593/abstract?

Nguyen, T. V. H., Ahmed, S., Chevapatrakul, T., & Onali, E. (2020). Do stress tests affect bank liquidity creation? *Journal of Corporate Finance*, *64*, 101622. https://doi.org/10.1016/j.jcorpfin.2020.101622

Nicoletti, A. (2018). The effects of bank regulators and external auditors on loan loss provisions. *Journal of Accounting and Economics*, *66*(1), 244–265. https://doi.org/10.1016/j.jacceco.2018.05.003

Pana, E., Park, J., & Query, T. (2010). The impact of bank mergers on liquidity creation. *Journal of Risk Management in Financial Institutions*, *4*(1).

Peng, L., & Xiong, W. (2006). Investor attention, overconfidence and category learning. *Journal of Financial Economics*, *80*(3), 563–602. https://doi.org/10.1016/j.jfineco.2005.05.003

Posner, E., & Weyl, E. G. (2013). Benefit-Cost Analysis for Financial Regulation. *American Economic Review*, *103*(3), 393–397. https://doi.org/10.1257/aer.103.3.393

Richards, N. (2013). Federal Deposit Insurance Corporation Improvement Act of 1991. *Federal Reserve Bank of Philidelphia.*

Richardson, S., Teoh, S. H., & Wysocki, P. D. (2004). The Walk-down to Beatable Analyst Forecasts: The Role of Equity Issuance and Insider Trading Incentives* [_eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1506/KHNW-PJYL-ADUB-0RP6]. *Contemporary Accounting Research*, *21*(4), 885–924. https://doi.org/10.1506/KHNW-PJYL-ADUB-0RP6

Sims, C. A. (2003). Implications of rational inattention. *Journal of Monetary Economics*, *50*(3), 665–690. https://doi.org/10.1016/S0304-3932(03)00029-1

So, E. C. (2013). A new approach to predicting analyst forecast errors: Do investors overweight analyst forecasts? *Journal of Financial Economics*, *108*(3), 615–640. https://doi.org/10.1016/j.jfineco.2013.02.002

Srivastav, A., & Vallascas, F. (2019). Is there a Benefit from Reduced Regulation on Small Banks? *SSRN Electronic Journal*. https://doi.org/10.2139/ssrn.3389946

Tran, V. T., Lin, C.-T., & Nguyen, H. (2016). Liquidity creation, regulatory capital, and bank profitability. *International Review of Financial Analysis*, *48*, 98–109. https://doi.org/10.1016/j.irfa.2016.09.010

Van Nieuwerburgh, S., & Veldkamp, L. (2009). Information Immobility and the Home Bias Puzzle. *The Journal of Finance*, *64*(3), 1187–1215. https://doi.org/10.1111/j.1540-6261.2009.01462.x

Van Nieuwerburgh, S., & Veldkamp, L. (2010). Information Acquisition and Under-Diversification. *The Review of Economic Studies*, *77*(2), 779–805. https://doi.org/10.1111/j.1467-937X.2009.00583.x

van Binsbergen, J., Han, X., & Lopez-Lira, A. (2020). Man vs. Machine Learning: The Term Structure of Earnings Expectations and Conditional Biases. *NBER*. https://doi.org/10.3386/w27843

van Binsbergen, J. H., Han, X., & Lopez-Lira, A. (2022). Man versus Machine Learning: The Term Structure of Earnings Expectations and Conditional Biases. *The Review of Financial Studies*, hhac085. https://doi.org/10.1093/rfs/hhac085

Veldkamp, L. (2011). *Information Choice in Macroeconomics and Finance*. Retrieved November 18, 2022, from https://press.princeton.edu/books/hardcover/9780691142203/information-choice-in-macroeconomics-and-finance

Veldkamp, L. L. (2006). Information Markets and the Comovement of Asset Prices. *The Review of Economic Studies*, *73*(3), 823–845. https://doi.org/10.1111/j.1467-937X.2006.00397.x

Zhang, X. F. (2006). Information Uncertainty and Stock Returns. *The Journal of Finance*, *61*(1), 105–137. https://doi.org/10.1111/j.1540-6261.2006.00831.x

Zheng, C., (Wai Kong) Cheung, A., & Cronje, T. (2019). The moderating role of capital on the relationship between bank liquidity creation and failure risk. *Journal of Banking & Finance*, *108*, 105651. https://doi.org/10.1016/j.jbankfin.2019.105651

# Appendix A

# Chapter 2 Appendix

## A.1 Banking Regulations

Financial institutions are subject to several other size-based thresholds. Table 2.1 shows the various thresholds for banks with assets under $20 billion as provided by Labonte and Perkins, 2021. Figure A.1 also provides graphical evidence of the distribution of banks above each threshold as of March 31, 2022. 93.99% of banks have assets above $50 million. The Interlock regulations, first effective October 1, 1996, prohibit management to work at multiple unaffiliated banks at one time (12 C.F.R §212 and 12 C.F.R §348). However, banks below the threshold are exempt from this, provided the two institutions are in the same MSA. The threshold was initially set at $20 million but was raised to $50 million as part of the Financial Services Regulatory Relief Act of 2006. Management at banks with assets above $10 billion are excluded from serving at any other institution with assets above $10 billion, regardless of the location. Initially, this threshold was that management at banks above $2.5 billion could not serve as management of unaffiliated banks with assets above $1.5 billion as of January 20, 2016 (80 FR 79252). This was adjusted to the $10 billion on October 10, 2019 (84 FR 54472). Banks above $50 million in assets have to comply with the reporting requirements of the Home Mortgage Disclosure Act (HMDA).[1] Bank Holding Companies with assets under $50 million are exempt from some activities related to selling insurance that are otherwise prohibited (12 C.F.R §225.28). This is part of Regulation Y and was implemented on April 21, 1997.

The streamlined SEC reporting allows small public banks to use the call report financial statements as substitutes for form 10-Q (12 C.F.R §208.36(b))[2]. This is part of Regulation H, which was implemented in 1952, but this section went into effect first on October 1, 1998 (63 FR 37646). As of December 27, 2002, a holding company that has assets above $150 million can apply for expedited processing of an application for deposit insurance (12 C.F.R §303.22). For bank holding companies with assets under $300 million, the requirements to pursue expedited acquisitions are less stringent (12 C.F.R §225.14 and 225.23). Banks below this threshold can also take on acquisitions above 35% of their assets as long as the total amount of the company after acquisition remains below $300 million. Above $3 billion, the pro-forma balance sheet requires more information. The $3 billion threshold has been adjusted following the size of what is considered a Small Holding Company. When this was first effective on April 21, 1997, this was $150 million (62 FR 9324). This was increased to $500 million effective March 30, 2006 (71 FR 9897), $1 billion effective May 15, 2015 (80 FR

[1]This number is adjusted annually and was introduced first with a threshold of $10 million
[2]This excludes banks that have foreign offices.

20153), and finally to $3 billion effective August 30, 2018 (83 FR 44199). Acquisitions are limited to an upper threshold of $7.5 billion.

Bank holding companies below $500 million are exempt from risk-based capital on a consolidated basis as long as certain requirements are met (12 C.F.R. §225 Appendix A). BHC below $500 million in assets may be exempt[3] if they:

1. Do not have significant non-banking activities. This is either directly or non-directly through a non-bank subsidiary.

2. Do not have significant off-balance sheet activities

3. Do not have material amounts of debt or equity securities outstanding that are registered with the SEC

Banks with assets under $1 billion in either of the two prior years as of December 31 are not required to establish flood insurance escrow accounts if the loans meet certain requirements related to flood zones (12 C.F.R. §339.5). The Truth in Lending Act (TILA) exempts banks with assets under $2.336 billion from the requirement to establish escrow accounts for some first-lien high-priced mortgages (12 C.F.R. §226). High-priced mortgages are defined as loans where the interest rate is at least 1.5 percentage points or 3.5 percentage points higher than the average prime offer rate[4] when established for first-lien and subordinate-lien loans, respectively.

[3]The regulation allows the Federal Reserve to apply the risk-based capital requirements if they deem it necessary.

[4]These are published by the FFEIC every week.

There are several regulations at the \$3 billion asset level. The Collins Amendment, section 171 of the Dodd-Frank Act in 2010, established general risk-based capital and leverage requirements for banks and holding companies. These regulations require that bank holding companies meet the capital requirements at the holding company level required by their bank subsidiaries. In addition, it provided an exemption from these rules for bank holding companies taking on greater levels of debt to complete an acquisition. These thresholds are based on the Small Bank Holding company thresholds previously discussed, so the threshold started at \$500 million and then was subsequently increased to \$1 billion and then \$3 billion. In 1974, the Federal Deposit Act established that banks with assets under \$500 million could be eligible for an 18-month instead of a 12-month examination cycle. In February 2015, this threshold was raised to \$1 billion (81 FR 10069). Finally, in August 2018, this was again raised to the current level of \$3 billion (83 FR 43965). The streamlined stock buyback and redemption reporting at \$3 billion are tied to the expedited acquisition activity threshold of the same amount previously discussed.

Bank holding company examinations are determined by the size of the holding company as well as its complexity. For holding companies identified as non-complex with assets below \$3 billion, if their examination currently is a 1 or 2, it is based on the rating of their subsidiaries as long as those subsidiaries have ratings of 1 or 2 (SR 13-21). For those non-complex holding companies with lower ratings or holding companies that are determined to be complex, an off-site review is conducted. For holding companies between \$3 billion and \$10 billion, non-complex holding companies with ratings of 1 or 2 are subject to targeted off-site inspections every two years. Those with lower ratings have full-scope off-site inspections annually. Complex holding companies have full-scope on-site inspections annually. For holding companies above $10 billion, the on-site inspections have more requirements (SR 17-12).

For banks above $5 billion in assets, they must file FFEIC 041 or 031. These are better known as Call Reports. Before the March 2017 Call Report, banks had the option to file FFEIC 041 or 031. Beginning for the quarter of March 30, 2017, banks with assets under $1 billion could file a simplified report under certain criteria. Banks with foreign offices, assets over $100 billion, or designated as an advanced approaches institution must file 031. Those not meeting these criteria can file 041. This includes institutions eligible for 051. This was set in place under the Economic Growth, Regulatory Relief, and Consumer Protection Act in 2018.

At $10 billion in assets, there are several regulatory thresholds, many of which are associated with the Dodd-Frank Act which was enacted on July 21, 2010. The act established the Bureau of Consumer Financial Protection and set it as the primary regulator for consumer compliance for all banks above $10 billion (Title X). The Durbin Amendment (section 1075) limited the fees that can be charged to retailers for processing debit cards. Dodd-Frank also allows for banks below the threshold to use alternative standards for Qualified Mortgages than those that were in place previously. This was an attempt to help give individuals access to credit that may have been denied due to situations such as too high of a Debt-to-Income ratio. Section 104 provides exemptions for requiring mortgage escrows for banks under the threshold as long as the firm meets other requirements. Section 201 required regulators to work to develop the Community Bank Leverage Ratio (CBLR). This allows banks that opt in to using the CBLR in replacement of the four regulatory capital ratios currently used. It does not use risk-weighted assets so requires less reporting as well. Dodd-Frank also established the requirement for non-cleared financial swap transactions to have a margin account. Initially, there was no threshold associated with this. This was changed to exempt banks below $10 billion under the Terrorism Risk Insurance Program Reauthorization Act of 2015.

The Volcker rule (section 619) established prohibitions for banks engaging in proprietary trading or having ownership interests in hedge funds or private equity funds. Initially, this applied to all banks. While the rule was put forth in the Dodd-Frank Act, the Volcker rule faced numerous delays in implementation. Updates for the Volcker Rule in addition to several other regulations came as part of the Economic Growth, Regulatory Relief, and Consumer Protection Act. On December 10, 2013, the final ruling was made for enactment as of April 1, 2014 (79 FR 5805). On November 11, 2019, an update to this rule was made to include thresholds for compliance (84 FR 62165). Banks below $1 billion assets are assumed to comply automatically (12 C.F.R §351). Banks below $10 billion in assets and trading assets and liabilities equal to five percent or less of total consolidated assets are exempt.

The Economic Growth, Regulatory Relief, and Consumer Protection Act also provided regulation to allow federal savings associations, sometimes known as thrifts to operate as a national bank, thus changing their regulatory regime without having to change their charter. This was available for thrifts that had assets below $20 billion as of December 31, 2017.

## A.2 Additional Tables

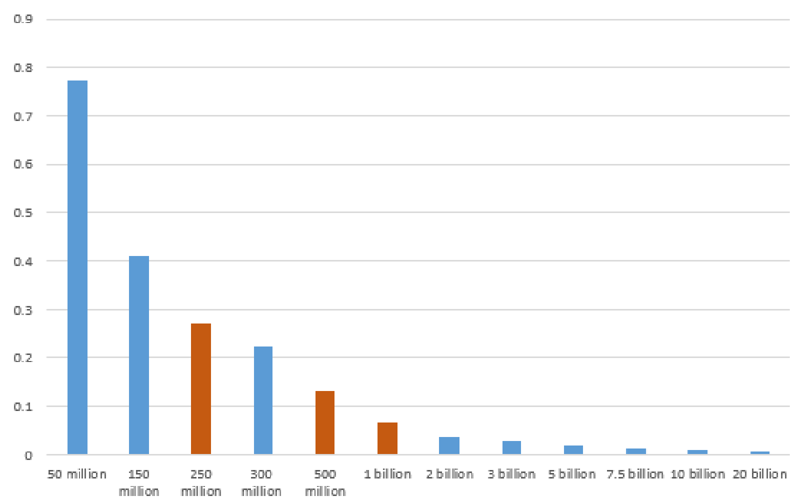



Figure A.1: Percent of Banks at various thresholds

This figure shows the percentage of the number of financial institutions as each threshold provided in Table 2.1 as of March 31, 2022. This only includes thresholds that have bank-specific regulatory size thresholds.

Table A.1: Summary Statistics: CRA

This table provides the summary statistics of the variables used in the analysis. This excludes firms required to comply with FDICIA All variables, excluding total assets, are winsorized at the 1% and 99% level (using the full sample).

| | N | Mean | SD | P25 | Median | P75 |
|---|---|---|---|---|---|---|
| Total Assets | 213126 | 126.764 | 102.969 | 54.738 | 93.188 | 159.898 |
| Non-Performing Loans | 213113 | 0.013 | 0.017 | 0.002 | 0.007 | 0.017 |
| ROA | 213126 | 0.009 | 0.008 | 0.006 | 0.010 | 0.014 |
| ROE | 213126 | 0.091 | 0.089 | 0.059 | 0.095 | 0.134 |
| Risk Based Cap. Ratio | 213126 | 0.171 | 0.073 | 0.123 | 0.148 | 0.192 |
| Liquidity Creation | | | | | | |
| LC(Total) | 213126 | 0.272 | 0.176 | 0.154 | 0.278 | 0.395 |
| LC(On) | 213126 | 0.223 | 0.155 | 0.122 | 0.232 | 0.334 |
| LC(Off) | 213126 | 0.048 | 0.033 | 0.024 | 0.041 | 0.065 |
| Liquid Assets | 213126 | 0.324 | 0.153 | 0.209 | 0.304 | 0.418 |
| Illiquid Assets | 213126 | 0.443 | 0.170 | 0.318 | 0.441 | 0.566 |
| Liquid Liabilities | 213126 | 0.440 | 0.123 | 0.355 | 0.432 | 0.516 |
| Illiquid Liabilities | 213126 | 0.112 | 0.032 | 0.090 | 0.104 | 0.126 |
| Liquid Derivatives | 213126 | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 |
| Liquid Guarantees | 213126 | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 |
| Illiquid Guarantees | 213126 | 0.096 | 0.067 | 0.048 | 0.083 | 0.130 |
| Small Business and Farm Lending | | | | | | |
| Number Business Loans | 53110 | 288.894 | 245.858 | 125.000 | 220.000 | 370.000 |
| Number Farm Loans | 53110 | 151.776 | 205.554 | 6.000 | 72.000 | 218.000 |
| Number All Loans | 53110 | 440.670 | 352.332 | 206.000 | 347.000 | 566.000 |
| Balance Outstanding Business Loans | 53110 | 23.471 | 22.105 | 7.782 | 16.450 | 31.964 |
| Balance Outstanding Farm Loans | 53110 | 7.404 | 10.103 | 0.407 | 3.475 | 10.501 |
| Balance Outstanding All Loans | 53110 | 30.875 | 24.152 | 14.039 | 24.148 | 40.643 |
| SBA Lending | | | | | | |
| Number Loans | 11058 | 3.109 | 5.870 | 1.000 | 1.000 | 3.000 |
| Loan Size | 11058 | 1.022 | 2.062 | 0.140 | 0.350 | 0.970 |
| Bank Portion | 11058 | 0.246 | 0.502 | 0.023 | 0.081 | 0.229 |
| Jobs Supported | 11058 | 29.165 | 79.518 | 0.000 | 0.000 | 19.000 |
| % Loan Failure | 10363 | 0.167 | 0.304 | 0.000 | 0.000 | 0.225 |
| Loan Quality | | | | | | |
| Past Due Loans 30-89 | 197017 | 0.015 | 0.014 | 0.005 | 0.011 | 0.021 |
| Past Due Loans 90+ | 213113 | 0.004 | 0.007 | 0.000 | 0.001 | 0.004 |
| Non-Accrual Loans | 213113 | 0.009 | 0.014 | 0.001 | 0.004 | 0.012 |
| Charge-Offs | 213113 | 0.004 | 0.007 | 0.001 | 0.002 | 0.005 |
| Recoveries | 213113 | 0.001 | 0.001 | 0.000 | 0.000 | 0.001 |
| Net Charge-Offs | 213113 | 0.003 | 0.007 | 0.000 | 0.001 | 0.004 |

Table A.2: Summary Statistics: FDICIA

This table provides the summary statistics of the variables used in the analysis. This excludes firms not required to comply with CRA All variables, excluding total assets, are winsorized at the 1% and 99% level (using the full sample).

| | N | Mean | SD | P25 | Median | P75 |
|---|---|---|---|---|---|---|
| Total Assets | 31656 | 497.907 | 187.259 | 347.492 | 439.320 | 620.680 |
| Non-Performing Loans | 31654 | 0.014 | 0.019 | 0.003 | 0.007 | 0.016 |
| ROA | 31656 | 0.009 | 0.008 | 0.007 | 0.010 | 0.014 |
| ROE | 31656 | 0.098 | 0.094 | 0.066 | 0.106 | 0.142 |
| Risk Based Cap. Ratio | 31656 | 0.150 | 0.061 | 0.114 | 0.131 | 0.162 |
| Liquidity Creation | | | | | | |
| LC(Total) | 31656 | 0.344 | 0.179 | 0.236 | 0.357 | 0.467 |
| LC(On) | 31656 | 0.273 | 0.155 | 0.184 | 0.288 | 0.381 |
| LC(Off) | 31656 | 0.070 | 0.038 | 0.042 | 0.064 | 0.091 |
| Liquid Assets | 31656 | 0.288 | 0.146 | 0.183 | 0.259 | 0.366 |
| Illiquid Assets | 31656 | 0.476 | 0.176 | 0.353 | 0.480 | 0.605 |
| Liquid Liabilities | 31656 | 0.466 | 0.131 | 0.377 | 0.458 | 0.547 |
| Illiquid Liabilities | 31656 | 0.106 | 0.029 | 0.088 | 0.099 | 0.115 |
| Liquid Derivatives | 31656 | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 |
| Liquid Guarantees | 31656 | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 |
| Illiquid Guarantees | 31656 | 0.139 | 0.076 | 0.085 | 0.128 | 0.182 |
| Small Business and Farm Lending | | | | | | |
| Number Business Loans | 7956 | 736.222 | 432.558 | 406.000 | 675.000 | 1,013.500 |
| Number Farm Loans | 7956 | 163.229 | 287.326 | 1.000 | 18.000 | 179.000 |
| Number All Loans | 7956 | 899.451 | 604.370 | 448.000 | 777.500 | 1,237.500 |
| Balance Outstanding Business Loans | 7956 | 74.138 | 37.402 | 46.457 | 71.359 | 99.822 |
| Balance Outstanding Farm Loans | 7956 | 9.076 | 15.280 | 0.023 | 1.389 | 10.722 |
| Balance Outstanding All Loans | 7956 | 83.214 | 41.658 | 53.038 | 80.615 | 110.867 |
| SBA Lending | | | | | | |
| Number Loans | 3363 | 5.748 | 12.026 | 1.000 | 2.000 | 6.000 |
| Loan Size | 3363 | 1.744 | 2.935 | 0.225 | 0.650 | 1.838 |
| Bank Portion | 3363 | 0.435 | 0.712 | 0.051 | 0.162 | 0.473 |
| Jobs Supported | 3363 | 57.906 | 114.023 | 0.000 | 10.000 | 60.000 |
| % Loan Failure | 3199 | 0.167 | 0.274 | 0.000 | 0.000 | 0.250 |
| Loan Quality | | | | | | |
| Past Due Loans 30-89 | 30799 | 0.011 | 0.011 | 0.004 | 0.008 | 0.015 |
| Past Due Loans 90+ | 31654 | 0.003 | 0.005 | 0.000 | 0.001 | 0.003 |
| Non-Accrual Loans | 31654 | 0.011 | 0.016 | 0.002 | 0.005 | 0.012 |
| Charge-Offs | 31654 | 0.005 | 0.008 | 0.001 | 0.002 | 0.005 |
| Recoveries | 31654 | 0.001 | 0.001 | 0.000 | 0.000 | 0.001 |
| Net Charge-Offs | 31654 | 0.004 | 0.007 | 0.000 | 0.001 | 0.004 |

Table A.3: Liquidity Creation Components: FDICIA

This table shows the results of the regulatory change to FDICIA ICFR requirements on the components of liquidity creation. *Post* is a dummy variable equal to 1 for years after 2005 and zero for years prior to 2005. *FDICIA* is an indicator variable equal to 1 if the bank had assets above $500 million at the beginning of the current year. Size is the natural log of Total Assets. Risk based capital ratio is equity/total risk weighted assets. ROE is the quarterly net income/Equity. Non-performing loans is the sum of non-accrual loans and loans past 90 days overdue scaled by total loans. The control variables are all lagged. Bank and state by year-quarter fixed effects are included and standard errors are double clustered at the year-quarter and state level. T-statistics are reported in the parentheses. The symbols *, **, and *** represent statistical significance at the 10%, 5%, and 1% levels, respectively.

| | (1)<br>Illiquid Assets | (2)<br>Illiquid Liabilities | (3)<br>Liquid Assets | (4)<br>Liquid Liabilities | (5)<br>Illiquid Guarantees |
|---|---|---|---|---|---|
| Post x FDICIA | 0.005 | 0.000 | -0.000 | 0.003 | 0.003 |
| | (1.03) | (0.34) | (-0.08) | (0.96) | (1.32) |
| Size | -0.023* | -0.005** | 0.056*** | -0.081*** | -0.013** |
| | (-1.76) | (-2.10) | (5.79) | (-11.95) | (-2.32) |
| Risk Based Cap. Ratio | -0.752*** | 0.288*** | 0.890*** | -0.003 | -0.254*** |
| | (-7.68) | (11.02) | (9.07) | (-0.06) | (-5.39) |
| ROA | 1.032*** | 0.416*** | -0.900*** | 0.525** | 1.228*** |
| | (4.48) | (7.03) | (-4.08) | (2.68) | (8.75) |
| Non-Performing Loans | -0.044 | -0.016 | 0.131 | -0.234*** | -0.459*** |
| | (-0.55) | (-1.08) | (1.50) | (-3.65) | (-10.31) |
| Constant | 0.719*** | 0.087*** | -0.178*** | 0.961*** | 0.252*** |
| | (8.58) | (5.48) | (-2.97) | (22.07) | (6.61) |
| Bank FE | Yes | Yes | Yes | Yes | Yes |
| State x Year-Quarter FE | Yes | Yes | Yes | Yes | Yes |
| Observations | 31505 | 31505 | 31505 | 31505 | 31505 |
| Adj. R2 | 0.951 | 0.870 | 0.917 | 0.912 | 0.857 |

Table A.4: Small Business Lending Non-Farm and C&I

This table shows the results of the regulatory change to CRA reporting on small business lending. Specifically this looks at non-farm and commercial and industrial loans. Column 1 looks at the number of non-farm loans outstanding, column 2 looks at the size of non-farm loans outstanding, column 3 looks at the number of C&I loans outstanding, and column 4 looks as the size of C&I loans outstanding. *Post* is a dummy variable equal to 1 for years after 2005 and zero for years prior to 2005. *CRA* is an indicator variable equal to 1 if the bank had assets above $250 million at the beginning of the current year and previous year. Size is the natural log of total assets. Risk based capital ratio is equity/total risk weighted assets. ROA is the quarterly net income/total assets. Non-performing loans is the sum of non-accrual loans and loans past 90 days overdue scaled by total loans. The control variables are all lagged. Bank and state by year fixed effects are included and standard errors are double clustered at the bank and year level. The symbols *, **, and *** represent statistical significance at the 10%, 5%, and 1% levels, respectively.

| | Non-Farm | | C&I | |
|---|---|---|---|---|
| | (1) | (2) | (3) | (4) |
| | No. Loans | $ Outstanding | No. Loans | $ Outstanding |
| Post x CRA | -0.862 | 3.645*** | -17.254 | 0.472 |
| | (-0.25) | (4.72) | (-1.51) | (0.82) |
| Size | 48.814*** | 10.679*** | 104.805*** | 7.553*** |
| | (13.96) | (17.34) | (9.21) | (14.25) |
| Risk Based Cap. Ratio | -77.324*** | -12.774*** | -170.431*** | -10.463*** |
| | (-6.01) | (-6.75) | (-7.15) | (-8.88) |
| ROA | 243.308*** | 24.666** | 779.088*** | 26.441** |
| | (3.67) | (2.94) | (4.34) | (2.92) |
| Non-Performing Loans | -29.719* | -7.091* | -150.241** | -17.985*** |
| | (-1.85) | (-1.92) | (-2.64) | (-4.68) |
| Constant | -118.079*** | -32.919*** | -254.126*** | -22.767*** |
| | (-7.28) | (-11.72) | (-4.98) | (-9.20) |
| Bank FE | Yes | Yes | Yes | Yes |
| State x Year FE | Yes | Yes | Yes | Yes |
| Observations | 52989 | 52989 | 52989 | 52989 |
| Adj. R2 | 0.831 | 0.884 | 0.796 | 0.859 |

Table A.5: Small Business Lending Farmland and Agricultural Production

This table shows the results of the regulatory change to CRA reporting on small farm lending. Specifically this looks at farmland and agricultural production loans. Column 1 looks at the number of loans outstanding secured by farmland, column 2 looks at the size of loans outstanding secured by farmland, column 3 looks at the number of agricultural production loans outstanding, and column 4 looks as the size of agricultural production loans outstanding. *Post* is a dummy variable equal to 1 for years after 2005 and zero for years prior to 2005. *CRA* is an indicator variable equal to 1 if the bank had assets above $250 million at the beginning of the current year and previous year. Size is the natural log of total assets. Risk based capital ratio is equity/total risk weighted assets. ROA is the quarterly net income/total assets. Non-performing loans is the sum of non-accrual loans and loans past 90 days overdue scaled by total loans. The control variables are all lagged. Bank and state by year fixed effects are included and standard errors are double clustered at the bank and year level. The symbols *, **, and *** represent statistical significance at the 10%, 5%, and 1% levels, respectively.

| | Farm Land | | Agricultural Production | |
|---|---|---|---|---|
| | (1) | (2) | (3) | (4) |
| | No. Loans | $ Outstanding | No. Loans | $ Outstanding |
| Post x CRA | 0.373 | 0.400** | -7.262 | -0.063 |
| | (0.23) | (2.34) | (-1.51) | (-0.39) |
| Size | 19.634*** | 1.970*** | 42.308*** | 1.843*** |
| | (5.91) | (5.02) | (5.07) | (4.40) |
| Risk Based Cap. Ratio | -10.758 | -1.543** | -29.829 | -2.285*** |
| | (-1.67) | (-2.79) | (-1.77) | (-3.76) |
| ROA | 85.149* | 11.067** | 292.273*** | 10.469** |
| | (2.12) | (3.13) | (3.35) | (2.98) |
| Non-Performing Loans | -32.738** | -5.671*** | -10.620 | -4.401*** |
| | (-2.70) | (-4.74) | (-0.42) | (-3.92) |
| Constant | -42.185** | -4.921** | -75.312* | -4.044* |
| | (-2.81) | (-2.78) | (-1.90) | (-2.12) |
| Bank FE | Yes | Yes | Yes | Yes |
| State x Year FE | Yes | Yes | Yes | Yes |
| Observations | 50890 | 50890 | 50890 | 50890 |
| Adj. R2 | 0.899 | 0.909 | 0.888 | 0.929 |

## Table A.6: Small Business Lending: FDICIA

This table shows the results of the regulatory change to FDICIA ICFR requirements on small business lending. Specifically this looks at the totals of non-farm and farm related loans as well as the overall small business lending. Column 1 looks at the number of small business (non farm) loans outstanding, column 2 looks at the size of small business (non farm) loans outstanding, column 3 looks at the number of agricultural loans outstanding, column 4 looks as the size of agricultural outstanding, column 5 looks at the overall number of small business loans outstanding, and column 6 looks at the overall size of small business loans outstanding. *Post* is a dummy variable equal to 1 for years after 2005 and zero for years prior to 2005. *FDICIA* is an indicator variable equal to 1 if the bank had assets above $500 million at the beginning of the current year. Size is the natural log of Total Assets. Risk based capital ratio is equity/total risk weighted assets. ROE is the quarterly net income/Equity. Non-performing loans is the sum of non-accrual loans and loans past 90 days overdue scaled by total loans. The control variables are all lagged. Bank and state by year fixed effects are included and standard errors are clustered at the year level. T-statistics are reported in the parentheses. The symbols *, **, and *** represent statistical significance at the 10%, 5%, and 1% levels, respectively.

| | Small Business | | Small Farm | | Total | |
|---|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (5) | (6) |
| | No. Loans | $ Outstanding | No. Loans | $ Outstanding | No. Loans | $ Outstanding |
| Post x FDICIA | -7.541 | 1.167 | 10.537 | 0.463 | 2.996 | 1.630 |
| | (-0.71) | (0.81) | (1.41) | (1.19) | (0.20) | (1.04) |
| Size | 314.691*** | 39.081*** | 56.830** | 2.936** | 371.521*** | 42.017*** |
| | (5.94) | (10.43) | (2.91) | (3.11) | (5.52) | (10.28) |
| Risk Based Cap. Ratio | -392.410 | -72.574*** | -39.990 | -7.711* | -432.400 | -80.285*** |
| | (-1.38) | (-4.44) | (-0.46) | (-2.11) | (-1.20) | (-4.75) |
| ROA | 1329.116* | 201.419*** | -78.906 | 15.829 | 1250.210* | 217.249*** |
| | (2.12) | (3.83) | (-0.55) | (1.56) | (1.82) | (4.10) |
| Non-Performing Loans | -444.837** | -79.987*** | -38.875 | -3.291 | -483.712* | -83.278*** |
| | (-2.55) | (-4.36) | (-0.49) | (-0.99) | (-2.03) | (-4.13) |
| Constant | -1.1e+03*** | -155.021*** | -177.785 | -7.864 | -1.3e+03** | -162.885*** |
| | (-3.29) | (-6.70) | (-1.41) | (-1.34) | (-2.98) | (-6.41) |
| Bank FE | Yes | Yes | Yes | Yes | Yes | Yes |
| State x Year FE | Yes | Yes | Yes | Yes | Yes | Yes |
| Observations | 7854 | 7854 | 7854 | 7854 | 7854 | 7854 |
| Adj. R2 | 0.903 | 0.903 | 0.972 | 0.971 | 0.934 | 0.912 |

Table A.7: SBA Lending: Difference in Groups

This table shows the differences in the results of the regulatory change to CRA reporting related to SBA guaranteed loans in CRA-eligible and non-CRA-eligible zip codes. *Post* is a dummy variable equal to 1 for years after 2005 and zero for years prior to 2005. The year 2005 is excluded as the regulation change was announced and implemented during that year. *CRA* is an indicator variable equal to 1 if the bank had assets above $250 million at the beginning of the current year and previous year. Size is the natural log of total assets. Risk based capital ratio is equity/total risk weighted assets. ROA is the quarterly net income/total assets. Non-performing loans is the sum of non-accrual loans and loans past 90 days overdue scaled by total loans. The control variables are lagged. Bank and state by year fixed effects are included and standard errors are clustered at the year level. The symbols *, **, and *** represent statistical significance at the 10%, 5%, and 1% levels, respectively.

| | Number Loans | Loan Size | Jobs Supported | % Loan Failure |
|---|---|---|---|---|
| Post x CRA | -0.880* | -0.303 | -30.372** | 0.051 |
| | (-1.89) | (-1.08) | (-2.50) | (0.62) |
| Size | -0.403 | -0.097 | -0.784 | -0.043 |
| | (-0.75) | (-0.34) | (-0.06) | (-0.57) |
| Risk Based Cap. Ratio | 5.584 | 2.875 | 96.096* | 0.025 |
| | (1.09) | (1.69) | (1.96) | (0.10) |
| ROA | -21.162 | -14.631 | -415.062 | 1.912 |
| | (-0.73) | (-1.72) | (-0.69) | (0.87) |
| Non-Performing Loans | 11.819 | -0.219 | 92.638 | -0.947 |
| | (1.23) | (-0.02) | (0.22) | (-0.59) |
| Constant | -4.181** | -2.196*** | -183.442*** | 0.241* |
| | (-2.80) | (-6.70) | (-14.93) | (2.03) |
| Bank FE | Yes | Yes | Yes | Yes |
| State x Year FE | Yes | Yes | Yes | Yes |
| Observations | 9996 | 9996 | 9996 | 9323 |
| Adj. R2 | 0.776 | 0.649 | 0.576 | 0.101 |

## Table A.8: SBA Lending: FDICIA

This table shows the results of the regulatory change to FDICIA ICFR requirements related to SBA guaranteed loans in CRA-eligible and non-CRA-eligible zip codes. *Post* is a dummy variable equal to 1 for years after 2005 and zero for years prior to 2005. The year 2005 is excluded as the regulation change was announced and implemented during that year. *FDICIA* is an indicator variable equal to 1 if the bank had assets above $500 million at the beginning of the current year. Size is the natural log of total assets. Risk based capital ratio is equity/total risk weighted assets. ROA is the quarterly net income/total assets. Non-performing loans is the sum of non-accrual loans and loans past 90 days overdue scaled by total loans. The control variables are lagged. Bank and state by year fixed effects are included and standard errors are clustered at the year level. The symbols *, **, and *** represent statistical significance at the 10%, 5%, and 1% levels, respectively.

| | CRA-Eligible | | | | Not CRA-Eligible | | | |
|---|---|---|---|---|---|---|---|---|
| | (1) Number Loans | (2) Loan Size | (3) Jobs Supported | (4) % Loan Failure | (5) Number Loans | (6) Loan Size | (7) Jobs Supported | (8) % Loan Failure |
| Post x FDICIA | 0.704 | 0.236 | 10.215 | -0.029 | -1.231 | -0.313 | -12.448 | 0.037*** |
| | (1.35) | (0.67) | (0.91) | (-0.50) | (-1.61) | (-1.11) | (-1.13) | (3.17) |
| Size | 0.991 | 0.231 | 67.146* | -0.477 | -0.011 | 0.065 | 30.236 | 0.039 |
| | (0.49) | (0.19) | (1.99) | (-1.63) | (-0.01) | (0.12) | (0.86) | (0.38) |
| Risk Based Cap. Ratio | 19.219 | 6.410 | 321.808 | -2.009 | 0.765 | -2.192 | -55.113 | 0.365 |
| | (1.73) | (1.80) | (1.65) | (-0.98) | (0.09) | (-0.51) | (-0.25) | (0.87) |
| ROA | -48.626 | -20.483 | -1.7e+03 | 4.589 | 31.526 | 17.090 | 100.303 | -1.161 |
| | (-1.11) | (-0.87) | (-1.53) | (0.78) | (0.52) | (0.97) | (0.09) | (-0.82) |
| Non-Performing Loans | 13.012 | 13.545 | 233.425 | -0.367 | -2.902 | -0.981 | 18.585 | 0.785 |
| | (0.64) | (0.86) | (0.36) | (-0.24) | (-0.09) | (-0.08) | (0.03) | (1.14) |
| Constant | -6.090 | -1.396 | -412.659* | 3.333 | 6.866 | 1.893 | -105.919 | -0.126 |
| | (-0.49) | (-0.19) | (-2.01) | (1.76) | (0.63) | (0.62) | (-0.49) | (-0.19) |
| Bank FE | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| State x Year FE | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Observations | 607 | 607 | 607 | 542 | 2301 | 2301 | 2301 | 2207 |
| Adj. R2 | 0.332 | 0.295 | 0.179 | 0.130 | 0.734 | 0.718 | 0.641 | 0.149 |

Table A.9: Past Due Loans: FDICIA

This table shows the results of the regulatory change to FDICIA ICFR requirements on past due loans. *Post* is a dummy variable equal to 1 for years after 2005 and zero for years prior to 2005. *FDICIA* is an indicator variable equal to 1 if the bank had assets above $500 million at the beginning of the current year. Size is the natural log of total assets. Risk based capital ratio is equity/total risk weighted assets. ROA is the quarterly net income/total assets. The control variables are all lagged. Bank and state by year fixed effects are included and standard errors are double clustered at the bank and year level. T-statistics are reported in the parentheses. The symbols *, **, and *** represent statistical significance at the 10%, 5%, and 1% levels, respectively.

| | (1)<br>Non-Performing Loans | (2)<br>Past Due Loans 30-89 | (3)<br>Past Due Loans 90+ | (4)<br>Non-Accrual Loans |
|---|---|---|---|---|
| Post x FDICIA | 0.001 | 0.000 | -0.000 | 0.002** |
| | (1.59) | (0.41) | (-1.26) | (2.09) |
| Size | -0.003 | 0.002 | 0.001 | -0.003 |
| | (-1.19) | (1.44) | (0.92) | (-1.47) |
| Risk Based Cap. Ratio | 0.033** | 0.013** | 0.008** | 0.025** |
| | (2.69) | (2.09) | (2.05) | (2.36) |
| ROA | -1.029*** | -0.207*** | -0.112*** | -0.896*** |
| | (-12.91) | (-8.63) | (-5.24) | (-12.44) |
| Constant | 0.034** | 0.002 | -0.001 | 0.032*** |
| | (2.50) | (0.25) | (-0.29) | (2.75) |
| Bank FE | Yes | Yes | Yes | Yes |
| State x Year-Quarter FE | Yes | Yes | Yes | Yes |
| Observations | 31503 | 30675 | 31503 | 31503 |
| Adj. R2 | 0.717 | 0.564 | 0.462 | 0.693 |

Table A.10: Charge Offs: FDICIA

This table shows the results of the regulatory change to FDICIA ICFR requirements on charge-offs and recoveries. *Post* is a dummy variable equal to 1 for years after 2005 and zero for years prior to 2005. *FDICIA* is an indicator variable equal to 1 if the bank had assets above $500 million at the beginning of the current year. Size is the natural log of total assets. Risk based capital ratio is equity/total risk weighted assets. ROA is the quarterly net income/total assets. The control variables are all lagged. Bank and state by year fixed effects are included and standard errors are double clustered at the bank and year level. T-statistics are reported in the parentheses. The symbols *, **, and *** represent statistical significance at the 10%, 5%, and 1% levels, respectively.

| | (1) Charge-Offs | (2) Recoveries | (3) Net Charge-Offs |
|---|---|---|---|
| Post x FDICIA | 0.001** | 0.000 | 0.000* |
| | (2.40) | (1.66) | (1.94) |
| Size | -0.002 | -0.000*** | -0.001 |
| | (-1.59) | (-2.75) | (-1.14) |
| Risk Based Cap. Ratio | 0.015*** | 0.002** | 0.012** |
| | (2.76) | (2.24) | (2.53) |
| ROA | -0.585*** | -0.019*** | -0.554*** |
| | (-12.42) | (-5.95) | (-12.59) |
| Constant | 0.017*** | 0.003*** | 0.013** |
| | (2.74) | (3.20) | (2.35) |
| Bank FE | Yes | Yes | Yes |
| State x Year-Quarter FE | Yes | Yes | Yes |
| Observations | 31503 | 31503 | 31503 |
| Adj. R2 | 0.741 | 0.537 | 0.732 |

Table A.11: Regulatory Costs

This table shows the results of the regulatory change to CRA reporting on regulatory compliance costs. *Post* is a dummy variable equal to 1 for years after 2005 and zero for years prior to 2005. *CRA* is an indicator variable equal to 1 if the bank had assets above $250 million at the beginning of the current year and previous year. Size is the natural log of Total Assets. Risk based capital ratio is equity/total risk weighted assets. ROA is the quarterly net income/total assets. The control variables are all lagged. Bank and state by year-quarter fixed effects are included and standard errors are double clustered at the bank and year-quarter level. The symbols *, **, and *** represent statistical significance at the 10%, 5%, and 1% levels, respectively.

| | (1) Pre-Tax ROA | (2) Loans Per Employee | (3) % Change Employees | (4) Average Salaries | (5) Average Salary per Employee | (6) Data Processing and Fixed Assets Exp. |
|---|---|---|---|---|---|---|
| Post x CRA | -0.001 | 0.106*** | 0.002** | -0.000 | 0.001* | -0.000** |
| | (-1.66) | (3.83) | (2.20) | (-1.27) | (2.00) | (-2.09) |
| Size | 0.004*** | 0.677*** | -0.018*** | -0.003*** | 0.004*** | -0.001*** |
| | (6.44) | (17.79) | (-12.81) | (-16.37) | (8.28) | (-15.70) |
| Risk Based Cap. Ratio | 0.023*** | -2.587*** | 0.023* | 0.000 | 0.002 | -0.002*** |
| | (6.31) | (-13.26) | (1.86) | (0.22) | (1.33) | (-5.40) |
| ROA | | 8.826*** | 0.622*** | -0.026*** | 0.027* | -0.033*** |
| | | (9.38) | (12.45) | (-4.42) | (1.96) | (-14.59) |
| Constant | -0.011*** | -0.607*** | 0.080*** | 0.030*** | 0.031*** | 0.013*** |
| | (-3.42) | (-3.45) | (11.91) | (35.20) | (13.49) | (30.27) |
| Bank FE | Yes | Yes | Yes | Yes | Yes | Yes |
| State x Year-Quarter FE | Yes | Yes | Yes | Yes | Yes | Yes |
| Observations | 213047 | 196919 | 213006 | 213047 | 213007 | 213047 |
| Adj. R2 | 0.601 | 0.900 | 0.037 | 0.809 | 0.839 | 0.801 |

Table A.12: Regulatory Costs: FDICIA

This table shows the results of the regulatory change to FDICIA ICFR requirements on regulatory compliance costs. *Post* is a dummy variable equal to 1 for years after 2005 and zero for years prior to 2005. *FDICIA* is an indicator variable equal to 1 if the bank had assets above $500 million at the beginning of the current year. Size is the natural log of Total Assets. Risk based capital ratio is equity/total risk weighted assets. ROA is the quarterly net income/total assets. The control variables are all lagged. Bank and state by year-quarter fixed effects are included and standard errors are double clustered at the bank and year-quarter level. The symbols *, **, and *** represent statistical significance at the 10%, 5%, and 1% levels, respectively.

| | (1) Pre-Tax ROA | (2) Loans Per Employee | (3) % Change Employees | (4) Average Salaries | (5) Average Salary per Employee | (6) Data Processing and Fixed Assets Exp. |
|---|---|---|---|---|---|---|
| Post x FDICIA | -0.001** | 0.036 | -0.002* | -0.000 | 0.001 | -0.000 |
| | (-2.07) | (1.30) | (-1.85) | (-0.59) | (1.19) | (-0.32) |
| Size | 0.008*** | 0.706*** | -0.012*** | -0.005*** | 0.002* | -0.002*** |
| | (3.62) | (7.88) | (-3.56) | (-9.44) | (1.70) | (-8.51) |
| Risk Based Cap. Ratio | 0.048*** | -2.347** | 0.013 | -0.004 | 0.004 | -0.001 |
| | (4.32) | (-2.54) | (0.78) | (-1.00) | (0.61) | (-0.51) |
| ROA | | 6.687*** | 0.778*** | 0.002 | 0.032 | -0.021*** |
| | | (4.83) | (10.13) | (0.30) | (1.39) | (-5.23) |
| Constant | -0.042*** | -1.457*** | 0.071*** | 0.046*** | 0.042*** | 0.019*** |
| | (-3.00) | (-2.89) | (3.53) | (13.88) | (6.17) | (11.71) |
| Bank FE | Yes | Yes | Yes | Yes | Yes | Yes |
| State x Year-Quarter FE | Yes | Yes | Yes | Yes | Yes | Yes |
| Observations | 31505 | 30658 | 31486 | 31505 | 31486 | 31505 |
| Adj. R2 | 0.681 | 0.933 | 0.054 | 0.888 | 0.896 | 0.872 |

# Appendix B

# Chapter 3 Appendix

## B.1 Input Data Construction

### B.1.1 Timing

We construct our train and test datasets carefully to ensure no data leakage. Figure B.1 gives an example of our forecast timeline. At the end of each month $t$, for each stock $i$, and for a specific forecast horizon $\tau$, we construct the earnings prediction. The target variable of interest is the analysts' one-quarter, two-quarter, one-year or two-year ahead forecasts error (i.e., the realized errors of analysts' forecasts made at month $t$).[1]

We train our models in rolling 10-year estimation windows. In our training set we make sure that both the target variable and the predictors in the train dataset are known at month $t$. Specifically, that means the realized earnings used in constructing the target variable are

[1]For analysts' forecasts from I/B/E/S, we use the consensus median forecasts as of the latest IBES statistical period, STATPERS, before the end of month $t$.

known/announced between $t - 120$ and $t$. After we fit the model, including selecting the optimal hyper-parameters, we use the fitted model to generate earnings predictions at month $t$. Hyper-parameter tuning is discussed below in B.2.

### B.1.2 Data Cleaning

When forecasting earnings for a specific forecast horizon $\tau$, we require the corresponding analysts' forecasts to be non-missing and the corresponding realized earnings to be unannounced as of month $t$. We delete about 0.2% of the one-year ahead observations, for which the I/B/E/S data are likely to contain errors if 1) the one-year ahead earnings as indicated by FPI=1 are already announced before the beginning of month $t$, or 2) the fiscal year ends of one-year ahead earnings are earlier than month $t - 6$ . We follow the same process for the two-year ahead (FPI=2), one-quarter ahead (FPI=6), and two-quarter ahead (FPI=7) earnings. The only variation is the cut-off period by which we delete observations likely to contain errors. For the two-year ahead earnings, we delete observations for which the fiscal year end is earlier than month $t + 6$. In addition, to account for the possible seasonality in the quarterly forecasts, we include dummy variables for the interaction of the firm's industry, as defined by the Fama-French 30, with the month of year in which the forecast is made.

For each 10-year estimation window ending at month $t$, we drop predictors that have more than 15% of their values missing in the estimation window. To alleviate the concerns of outliers influencing the model estimation, we winsorize all input variables at the 1% and 99% levels in each estimation window and then fill missing predictor values using the variable's median value.[2] We repeat this process for the test set as well. In our training dataset, we trim the analysts' forecast errors at the 1% and 99% levels. Then we standardize the predictor values in the training dataset prior to fitting our model. We use the same mean and standard deviation to standardize our predictors for our testing set. After our predictions are generated, we drop any firm-month combinations that do not have either the one-year-ahead or two-year-ahead analysts' forecasts. We also drop observations where either the absolute value of the analysts' forecast scaled by price or our machine forecast scaled by price are greater than 1 (this drops less than 500 observations).

[2]The only exception is the Fama French 30 Industry variable, in which we fill missing values using 30 to indicate that the industry is undefined.

Table B.1: WRDS Financial Ratio Variables

| Acronym | Definition | Acronym | Definition |
|---|---|---|---|
| accrual | Accruals/Average Assets | inv_turn | Inventory Turnover |
| adv_sale | Advertising Expenses/Sales | invt_act | Inventory/Current Assets |
| aftret_eq | After-tax Return on Average Common Equity | lt_debt | Long-term Debt/Total Liabilities |
| aftret_equity | After-tax Return on Total Stockholders Equity | lt_ppent | Total Liabilities/Total Tangible Assets |
| aftret_invcapx | After-tax Return on Invested Capital | npm | Net Profit Margin |
| at_turn | Asset turnover | ocf_lct | Operating Cash Flow/Current Liabilities |
| bm | Book/Market | opmad | Operating Profit Margin After Depreciation |
| capei | Shillers Cyclically Adjusted P/E Ratio | opmbd | Operating Profit Margin Before Depreciation |
| caital_ratio | Capitalization Ratio | pay_turn | Payables Turnover |
| cash_conversion | Cash Conversion Cycle (Days) | pcf | Price/Cash Flow |
| cash_debt | Cash Flow/Total Debt | pe_exi | P/E (Diluted, Excl. EI) |
| cash_lt | Cash Balance/Total Liabilities | pe_inc | P/E (Diluted, Incl. EI) |
| cash_ratio | Cash Ratio | pe_op_basic | Price/Operating Earnings (Basic, Excl. EI) |
| cfm | Cash Flow Margin | pe_op_dil | Price/Operating Earnings (Diluted Excl. EI) |
| curr_debt | Current Liabilities/Total Liabilities | pretret_earnat | Pre-tax Return on Total Earning Assets |
| curr_ratio | Current Ratio | pretret_noa | Pre-tax Return on Net Operating Assets |
| de_ratio | Total Debt/Total Equity | profit_lct | Profit Before Depreciation/Current Liabilities |
| debt_assets | Total Debt/TotalAssets | ps | Price/Sales |
| debt_at | Total Debt/TotalAssets | ptb | Price/Book |
| debt_capital | Total Debt/Total Capital | ptpm | Pre-Tax Profit margin |
| debt_ebitda | Total Debt/EBITDA | quick_ratio | Quick Ratio |
| debt_invcap | Long-term Debt/Invested Capital | rd_sale | Research and Development/Sales |
| divyield | Dividend Yield | rect_act | Receivables/Current Assets |
| dltt_be | Long-term Debt/Book Equity | rect_turn | Receivables Turnover |
| equity_invcap | Common Equity/Invested Capital | roa | Return on Assets |
| evm | Enterprise Value Multiple | roce | Return on Capital Employed |
| gpm | Gross Profit Margin | roe | Return on Equity |
| gprof | Gross Profit/Total Assets | sale_equity | Sales/Stockholders Equity |
| int_debt | Interest/Average Long-term Debt | sale_invcap | Sales/Invested Capital |
| int_totdebt | Interest/Average Total Debt | short_debt | Short-Term Debt/Total Debt |
| intcov | After-tax Interest Coverage | totdebt_invcap | Total Debt/Invested Capital |
| intcov_ratio | Interest Coverage Ratio | | |

Table B.2: Other Input Variables

| Acronym | Definition |
|---|---|
| dist1 | Time to FY1 Forecast Period End Date |
| dist2 | Time to FY2 Forecast Period End Date |
| dist6 | Time to FQ1 Forecast Period End Date |
| dist7 | Time to FQ2 Forecast Period End Date |
| ibes_earnings_ann | Most Recent Realized Annual Earnings |
| ibes_earnings_qtr | Most Recent Realized Quarterly Earnings |
| ffi30 | Fama-French Industry (30) |
| prc | Price |
| ret | One Month Return |
| size | LN(Market Cap) |
| datadate_F1ana_fe_med | Most Recent Realized FY1 Analyst Forecast Error |
| datadate_F2ana_fe_med | Most Recent Realized FY2 Analyst Forecast Error |
| qdate_Fqtr6ana_fe_med | Most Recent Realized FQ1 Analyst Forecast Error |
| qdate_Fqtr7ana_fe_med | Most Recent Realized FQ2 Analyst Forecast Error |
| medest1 | Most Recent FY1 Consensus Analyst Forecast |
| medest2 | Most Recent FY2 Consensus Analyst Forecast |
| medest6 | Most Recent FQ1 Consensus Analyst Forecast |
| medest7 | Most Recent FQ2 Consensus Analyst Forecast |
| rev_fpedats1_3m | Most Recent FY1 Analyst 3-month Revision |
| rev_fpedats2_3m | Most Recent FY2 Analyst 3-month Revision |
| rev_fpedats6_3m | Most Recent FQ1 Analyst 3-month Revision |
| rev_fpedats7_3m | Most Recent FQ2 Analyst 3-month Revision |

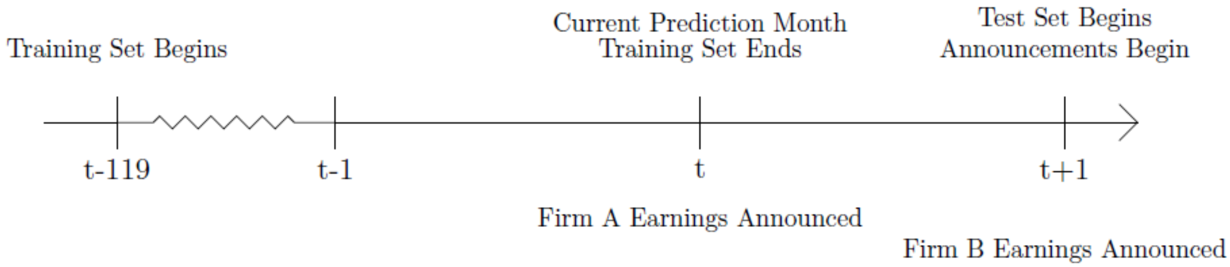



Figure B.1: Forecast Timeline

This figure gives an example of the timeline in our analysis. The training set consists of firms with earnings announcement dates between month *t-119* and *t*. The test sets then consists of firms with analysts' forecasts available in month *t* and announcement dates after month *t*. Observations are sorted into forecast periods based on their I/B/E/S FPI variable. In this figure Firm A would be part of the training set because its earnings have been announced as of month *t*. Firm B however would be part of the test set because its earnings have yet to be announced as of time *t*.

## B.2 Machine Learning Methodology

**Machine Learning Algorithms** This section gives an overview of our machine learning algorithms. All of our models use the same predictor sets as described in B.1. Our main specification uses the non-linear predictions generated by a gradient boosting decision tree algorithm as J. van Binsbergen et al., 2020 has shown that tree-based non-linear models outperform linear models. Other models are discussed briefly in SectionB.2.1.

**Gradient Boosting Decision Trees** Gradient boosted decision tree (GBDT) models are non-linear non-parametric ensemble models which combine the predictions of many decision trees. There are a number of variations on GBDT, but we utilize LightGBM, developed by Microsoft. Trees are grown in an adaptive way to correct the prediction error from the previous iteration, which is known as boosting (Friedman, 2001). The weighted average of these individual tree models is the final predictor. As Friedman, 2002 shows, subsampling helps reduce the computation time and the overfitting risk in boosting. Instead of randomly selecting a fraction of the data to train the model, we use Gradient-based One-Side Sampling (GOSS) (Ke et al., 2017) to sample observations. At each iteration, GOSS keeps data instances with residual errors in the top $a$ percentile and randomly selects $b$ percent of the remaining instances.[3] GOSS then combines these selected data instances to grow the next tree.

[3]The randomly selected data are amplified by the ratio of $\frac{1-a}{b}$ to minimize the influence on the data distribution.

Because the GBDT model grows iteratively (each consecutive tree fits the residual errors from the previous model) it is more prone to over-fitting than other models like random forest. One way the literature reduces the over-fitting risk is by using weaker learners, i.e., trees with a depth of 1 or 2 (e.g., Gu et al., 2020). We however allow our trees to be much deeper, setting our max depth at 7 to allow for more complex interactions between variables within each tree. However, we then limit the size of the tree by setting the number of leaves on each tree to 31 to help reduce over-fitting. This allows the trees to have complex interactions between variables, while still keeping the trees relatively simple. The other way the literature reduces over-fitting in GBDT models is by using a small learning rate. A small learning rate means more each individual tree corrects for a smaller amount of the residual errors. This leads to slower convergence but also decreases the risk of converging to a sub-optimal solution. We set the number of trees to be 2000 and use the cross validation to find the optimal learning rate between the range of [0.0001, 0.2]. To reduce the risk-of overfitting, once we identify our tuned hyper-parameters, we increase the number of trees to 4000 when training the model and implement early stopping. Early stopping will stop the model from generating further trees after the model's validation score[4] does not improve for a given number of subsequent trees. We set the early stopping cutoff to 10 consecutive trees.

**Cross-Validation and Hyper-Parameter Tuning** Because hyper-parameter tuning is computationally intensive, we tune the hyper-parameters once a year in June with the pre-

[4]We create our validation sample by randomly selecting 10% of the observations in the training set to be set aside. We train our model on the remaining 90% of the training dataset and use the validation set for the early stopping.

vailing data in the 10-year rolling window for each of our algorithms.[5] Our hyper-parameters are selected using a five-fold cross-validation. Similar to Kozak et al., 2020, our $K$-fold cross-validation splits the sample in the estimation window equally (time-series wise) into a $K$ sub-samples, so that each of the folds are non-overlapping in time. For a given combination of hyper-parameters, we withhold each of the $K$ subsamples, treat it as the out-of-sample data, and use the remaining data to fit the model. We then use the withheld subsample to compute the value of the loss function for the fitted model. This is repeated $K$ times. The cross-validated score is the average value of the loss function generated from the $K$ withheld subsamples for a given set of hyper-parameters. The hyper-parameters associated with the best cross-validation score are then selected.[6]

We use the optimal hyper-parameters estimated in June of year $t$ for fitting the models between June of year $t$ to May of year $t+1$. When fitting the model at the end of month $t$, we use all the prevailing data from the 10-year rolling estimation window ending at month $t$. Finally, we apply the fitted model to the predictor values in month $t$ to generate our predictions. The accuracy of these predictions is then evaluated using the realized values, which occur after month $t$ (i.e., the test sample).

[5] Gu et al., 2020 also tune their hyper-parameters once a year, while J. van Binsbergen et al., 2020 tune their hyper-parameters using the data up to 1985 and keep them fixed afterwards.

[6] Since it is computationally infeasible to evaluate all potential values in the parameter space, we use the Bayesian optimization algorithm to find the optimal combination of hyper-parameters for our GBDT algorithm (e.g., see a tutorial here Frazier, 2018). Specifically, we use skopt.BayesSearchCV to tune the hyper-parameters, which is a Bayesian search algorithm that samples a fixed number of combinations from the (approximated) Gaussian posterior distribution of the loss function. For each hyper-parameter tuned, we increased the number of combinations tried by 20. We use the log-uniform distribution for priors of all continuous hyper-parameters. The Bayesian search keeps track of the cross validation scores of already attempted parameter combinations and uses the results to find the most promising sets of parameters to try next.

### B.2.1 Implementation of Other Algorithms

We implement our algorithms using Python packages. We provide details of the other algorithms used as robustness tests. For the ease of replicating our results, we list the packages and the key hyper-parameters. All unlisted hyper-parameters are their default values.

#### Random Forest

Random forest is non-linear non-parametric ensemble model that combines the predictions of many decision trees in order to achieve variance reduction by averaging over de-correlated trees. We use LightGBM to run our random forest (RF) models. We train three hyper-parameters. First, random forest uses bootstrapped samples of the training set to build individual tree models (i.e., bagging as in Breiman, 1996). We control the fraction of observations sampled using the subsample parameter. Second, random forest grows the decision tree model by randomly selecting a percentage of the predictors to consider in each tree. We control this fraction using the col_sample_by_tree parameter. Finally, random forest grows trees until either a maximum number of leaves, or a maximum tree depth is achieved. We leave the maximum number of leaves at its default value, and instead tune the maximum tree depth to prevent overfitting. These are based on the parameters suggested byGu et al., 2020 and J. van Binsbergen et al., 2020.

#### Neural Networks

Artificial Neural Networks make predictions using a process inspired by the neurons of the human brain. Neural networks can be thought of as a generalization of the linear model, and they work by using multiple stages. First, the inputs are used to create hidden units, which are linear combinations of the various inputs. Then the output is predicted using a combination of the hidden units in the second stage. This second stage can be non-linear, which is why neural networks are a generalization of the linear model. More than two stages can be used, with intermediate stages adding an additional layer of hidden traits where each layer's hidden traits are a function of the hidden traits in the previous layer. At each layer, a number of nodes are used to process the data. Each node is an individual linear regression. However, the weights used in the regression are different at each node. We utilize a "feed forward" neural network which means that the results from one layer are passed forward to the next layer.

We base our neural networks on the methodology presented in Gu et al., 2020. We tune the L1 regularization penalty, which controls the weight on the L1 parameter penalization used at each hidden layer, and the learning rate, which controls the initial rate at which the optimizer converges[7]. We also tune the number of layers to allow for time varying model complexity. Our first hidden layer has 32 nodes, and each subsequent layer has half as many nodes as the previous layer. We follow the process in Gu et al., 2020 which uses the most recent set of data in the training set as the validation set in order to preserve temporal ordinality. We use only the most recent fold from our cross validation scheme as the validation set when training our models. [8]

[7]We use the Adam optimizer, which is an efficient implementation of the learning rate shrinkage algorithm

[8]We implement our neural network using Tensorflow.Keras. Unlike with our other models, we tune out hyper-parameters using Tensorflow.Keras.Keras_Tuner.BayesianOptimization, which is a Bayesian search algorithm similar to skopt.BayesSearchCV, but it is more compatible with Tensorflow.Keras. We use 50 trials with 5 initial points to tune the number of layers, the size of the L1 regularizer, and the learning rate.

Table B.3: Key Hyper-Parameter Values

This table shows the values we set for the various hyper-parameters in each python package. Any hyper-parameters not included here are set to their default values. If the variable name is missing, then the parameter is implemented manually.

Panel A: LightGBM

| Hyper-Parameter | Variable Name | Value/Range |
|---|---|---|
| Learning Rate | learning_rate | $\mathbb{R} \in [0.001, 0.2]$ |
| Number of Trees | n_estimators | Training: 2000, Testing: 4000 |
| Decision Tree Max Depth | max_depth | 7 |
| Model Type | boosting_type | GOSS |
| Minimum Observations Per Leaf | min_child_samples | 5 |

Panel B: Random Forest

| Hyper-Parameter | Variable Name | Value/Range |
|---|---|---|
| Subsample | max_samples | $\mathbb{R} \in [0.001, 0.5]$ |
| Feature Fraction | max_features | $\mathbb{R} \in [0.001, 0.999]$ |
| Number of Trees | n_estimators | 500 |
| Decision Tree Max Depth | max_depth | 4-10 |
| Minimum Observations Per Leaf | min_samples_leaf | 5 |

Panel C: Neural Network

| Hyper-Parameter | Variable Name | Value/Range |
|---|---|---|
| Learning Rate | learning_rate | $\mathbb{R} \in [0.0001, 0.01]$ |
| L1 Penalization | regularizers.L1 | $\mathbb{R} \in [1e-7, 1e-3]$ |
| Hidden Layers | | [1,2,3,4,5] |
| Activation Function | activation | ReLU |
| Epochs | epochs | 100 |
| Batch Size | batch_size | 10000 |
| Early Stopping Patience | patience | 5 |
| Early Stopping Metric | monitor | val_loss |
| Optimizer | optimizer | Adam |
| Ensemble | | 10 |

## B.3 The Effects of Uncertainty in Canonical Information Choice Models

The existing literature has emphasized the positive effect of uncertainty on attention and rationality of stock prices.

The intuition is that when the uncertainty level $(\frac{1}{\tau_{AF}^{-1}}+\frac{1}{\tau_0^{-1}})$ increases, information becomes more valuable, and investors optimally allocate more attention (higher $C^*$), which results in higher precision of their prediction signal (higher $\tau_s^*$), more debiasing (lower $b^*$), and thus more more rational investors' earnings expectations.

Since the optimal information choice is determined by the marginal benefit and cost of doing so, we first explain why the existing literature seems to focus exclusively on the effect of uncertainty on information benefit. Note that when the information processing capacity constraint is binding, investors' expected loss is

$$\frac{r}{2}\frac{1}{\tau_s+\tau_0}+\frac{\lambda}{2}ln\left(\frac{\tau_s+\tau_0}{\tau_{AF}+\tau_0}\right)$$

The marginal benefit of information acquisition (i.e., increasing $\tau_s$) is $-\frac{r}{2}\left(\tau_s+\tau_0\right)^{-2}$ and the marginal cost is $\frac{\lambda}{2}\left(\tau_s+\tau_0\right)^{-1}$. At the optimal choice ($\tau_s^*=\frac{r}{\lambda}-\tau_0$), the marginal benefit is equal to the marginal cost

$$\frac{r}{2}\left(\tau_s^* + \tau_0\right)^{-2} = \frac{\lambda}{2}\left(\tau_s^* + \tau_0\right)^{-1} = \frac{\lambda^2}{2r}$$

Now suppose the uncertainty increases, that is, $\tau_0^{'} = \tau_0 - \Delta$. Both the marginal benefit and the marginal cost increase. The former is because one additional unit of information can reduce more uncertainty and thus more valuable, while the latter is because under the entropy-based information constraint it is more costly to learn about the risk when little is understood.

Nevertheless, if $\lambda$ does not vary with $\tau_0$ as assumed in the existing literature, the increase in uncertainty will increase the marginal benefit more than the marginal cost at the previous optimal choice, following how the loss function in the objective is set up and the entropy-based learning rule:

$$\begin{aligned} &\frac{r}{2}\left(\tau_s^* + \tau_0^{'}\right)^{-2} - \frac{\lambda}{2}\left(\tau_s^* + \tau_0^{'}\right)^{-1} \\ =&\frac{\left(\frac{r}{\lambda} - \Delta\right)^{-2}}{2\lambda}\Delta > 0 \end{aligned}$$

As a result, there is an unambiguous positive relation between uncertainty and the rationality of market expectations ($\frac{\partial \tau_s^*}{\partial \tau_0^{-1}} > 0$).

**Solutions of Eq. (3.1)**

Following Baley and Veldkamp, 2021, we solve the model in two steps. First, given $s$, the optimal earnings forecast for investors is

$$a^* = E\left(y_{i,t+1}|s\right) = \frac{\tau_s}{\tau_0 + \tau_s} s \tag{B.1}$$

The resulting expected loss is $E\left[(y-a)^2\right] = \frac{1}{\tau_0+\tau_s}$, which decreases as $\tau_s$ increases. Thus, investors will unravel the bias in order to reduce entropy until the information processing capacity constraint binds, i.e., $\frac{\tau_s+\tau_0}{\tau_{AF}+\tau_0} = e^{2\kappa}$. Plug the optimal forecast in Eq. (B.1) and the binding capacity constraint into Eq. (3.1), we have

$$\min_\kappa \frac{r}{2} \frac{1}{(\tau_{AF}+\tau_0)e^{2\kappa}} + \lambda\kappa$$

Take the first order condition with respect to $\kappa$, we have

$$\frac{-re^{-2\kappa^*}}{(\tau_{AF}+\tau_0)} + \lambda = 0$$

We can then solve for $\kappa^*$,$C^*$, and $\tau_s^*$:

$$\kappa^* = \frac{1}{2} ln\left(\frac{r}{\lambda} \times (\tau_{AF} + \tau_0)^{-1}\right) \tag{B.2}$$

$$C^* = \frac{\lambda}{2} ln\left(\frac{r}{\lambda} \times (\tau_{AF} + \tau_0)^{-1}\right) \tag{B.3}$$

$$\tau_s^* = \frac{r}{\lambda} - \tau_0 \tag{B.4}$$

The extent of debiasing under the optimal solution ($b^*$) is as follows:

$$\frac{\tau_s^* + \tau_0}{\tau_{AF} + \tau_0} = \frac{1}{\frac{\lambda}{r}(\tau_{AF} + \tau_0)} \tag{B.5}$$

$$\tau_s^* + \tau_0 = \frac{r}{\lambda} \tag{B.6}$$

$$\tau_\eta^{-1} + \tau_B^{-1} \times (1 - b^*)^2 = \left(\frac{r}{\lambda} - \tau_0\right)^{-1} \tag{B.7}$$

$$(1 - b^*)^2 = \frac{\left(\frac{r}{\lambda} - \tau_0\right)^{-1} - \tau_\eta^{-1}}{\tau_B^{-1}} \tag{B.8}$$

$$(1 - b^*)^2 = \frac{\left(\frac{r}{\lambda} - \tau_0\right)^{-1} - \tau_\eta^{-1}}{\tau_B^{-1}} \tag{B.9}$$

When $\frac{r}{\lambda} > \bar{c} = \tau_\eta + \tau_0$, the information constraint in Eq. (3.2) is not binding. In this case, we have the following optimal solutions:

$$C^* = \frac{\lambda}{2} log \frac{\tau_\eta + \tau_0}{\tau_{AF} + \tau_0}, \tag{B.10}$$

$$\kappa^* = \frac{1}{2} log \frac{\tau_\eta + \tau_0}{\tau_{AF} + \tau_0} \tag{B.11}$$

$$\tau_s^* = \tau_\eta \tag{B.12}$$

$$b^* = 1 \tag{B.13}$$

**Proof of Prediction 1**

From Eqs. (3.4) and (3.6),

$$\frac{\partial C^*}{\partial r} = \frac{\lambda}{2r} > 0$$
$$\frac{\partial \tau_s^*}{\partial r} = \frac{1}{\lambda} > 0.$$

Since $\frac{\partial \tau_0}{\partial \lambda} > 0$,

$$\frac{\partial \tau_s^*}{\partial \lambda} = -\frac{r}{\lambda^2} - \frac{\partial \tau_0}{\partial \lambda} > 0$$

When $\frac{r}{\lambda}$ is sufficiently high, investors' information constraint is no binding. This is because from Eq. (3.2),

the information capacity needed to fully unravel analysts' bias is

$$\overline{\kappa} = ln \frac{\tau_\eta + \tau_0}{\tau_{AF} + \tau_0}$$

So if

$$ln\left(\frac{r}{\lambda}\right) - ln\left(\tau_{AF} + \tau_0\right) > \overline{\kappa}$$

$$ln\left(\frac{r}{\lambda}\right) > ln\left(\tau_\eta + \tau_0\right)$$

$$\frac{r}{\lambda} > \overline{c} = \tau_\eta + \tau_0$$

investors' unravel analysts' bias completely $b^* = 1$, $\tau_s^* = \tau_\eta$, and $\kappa^* = \frac{1}{2} log \frac{\tau_\eta + \tau_0}{\tau_{AF} + \tau_0}$.

**Proof of Prediction 2**

From Eq. (3.6), if $\frac{\partial \lambda}{\partial \tau_0^{-1}} = 0$

$$\frac{\partial \tau_s^*}{\partial \tau_0^{-1}} = \tau_0^2 > 0;$$

but if $\frac{\partial \lambda}{\partial \tau_0^{-1}} > 0$,

$$\frac{\partial \tau_s^*}{\partial \tau_0^{-1}} = \tau_0^2 - \frac{r}{\lambda^2} \frac{\partial \lambda}{\partial \tau_0^{-1}}$$

Therefore, if $\frac{\partial \lambda}{\partial \tau_0^{-1}}$ is sufficiently positive, $\frac{\partial \tau_s^*}{\partial \tau_0^{-1}} < 0$.

Note that the rationality of earnings expectations is measured by $b^*$ in our model and the partial derivatives of $b^*$ with respect to $r$, $\lambda$,and $\tau_0^{-1}$ have the same signs as the corresponding derivatives of $\tau_s^*$.

From Eq. (3.7),

$$\frac{\partial b^*}{\partial r} = \frac{\partial b^*}{\partial \tau_s^*} \times \frac{\partial \tau_s^*}{\partial r}$$
$$\frac{\partial b^*}{\partial \lambda} = \frac{\partial b^*}{\partial \tau_s^*} \times \frac{\partial \tau_s^*}{\partial \lambda}$$
$$\frac{\partial b^*}{\partial \tau_0^{-1}} = \frac{\partial b^*}{\partial \tau_s^*} \times \frac{\partial \tau_s^*}{\partial \tau_0^{-1}}$$

$$\begin{aligned}\frac{\partial b^*}{\partial \tau_s^*} &= -\left(\frac{\tau_s^{*-1} - \tau_\eta^{-1}}{\tau_B^{-1}}\right)^{-\frac{1}{2}} \times -\frac{\tau_s^{*-2}}{\tau_B^{-1}} \\ &= \left(\frac{\tau_B^{-\frac{1}{2}} \tau_s^{*-2}}{\sqrt{\tau_s^{*-1} - \tau_\eta^{-1}}}\right) > 0\end{aligned}$$

**Proof of Proposition 4**

$$\begin{aligned}C^* &= -\frac{\lambda}{2} ln\left(\frac{\lambda}{r} \times (\tau_{AF} + \tau_0)\right) \\ &= -\frac{\lambda}{2}\left[ln\left(\frac{\lambda}{r}\right) + ln\left(\frac{1}{\tau_{AF}^{-1}} + \frac{1}{\tau_0^{-1}}\right)\right]\end{aligned}$$

If $\frac{\partial \lambda}{\partial \tau_0^{-1}} = 0$

$$\frac{\partial C^*}{\partial \tau_0^{-1}} = \frac{\lambda}{2} \frac{\tau_0^2}{(\tau_{AF} + \tau_0)}$$
$$= \frac{\lambda}{2} \frac{1}{\left(\tau_{AF}\tau_0^{-2} + \tau_0^{-1}\right)}$$

and when the uncertainty changes are temporary time-series changes (denoted by superscript $TS$), and the partial derivative is evaluated at high fundamental uncertainty (e.g. high moving average 36-month IVOL), i.e. $\tau_0^{-1} = \tau_{0,h}^{-1}$, we have

$$\frac{\partial C^*}{\partial \tau_0^{-1}}|^{TS}_{\tau_{0,h}^{-1}} = \frac{\lambda}{2} \frac{1}{\left(\tau_{AF}\tau_{0,h}^{-2} + \tau_{0,h}^{-1}\right)}$$

and the increase in attention allocation should be lower for firms with higher fundamental uncertainty, i.e.

$$\frac{\partial C^*}{\partial \tau_0^{-1}}|^{TS}_{\tau_{0,h}^{-1}} < \frac{\partial C^*}{\partial \tau_0^{-1}}|^{TS}_{\tau_{0,l}^{-1}}$$

In the presence of the information cost channel (i.e., $\frac{\partial \lambda}{\partial \tau_0^{-1}} > 0$), we have

$$\frac{\partial C^*}{\partial \tau_0^{-1}}|^{TS} = \frac{1}{2}\left[ln\left(\frac{r}{\lambda}\right) - ln\left(\tau_{AF} + \tau_0\right)\right] \frac{\partial \lambda}{\partial \tau_0^{-1}}|^{TS} - \frac{\lambda}{2}\left[\frac{1}{\lambda}\frac{\partial \lambda}{\partial \tau_0^{-1}} - \frac{\tau_0^2}{\left(\tau_{AF} + \frac{1}{\tau_0^{-1}}\right)}\right]$$
$$= \frac{1}{2}\left[ln\left(\frac{r}{\lambda}\right) - ln\left(\tau_{AF} + \tau_0\right) - 1\right] \frac{\partial \lambda}{\partial \tau_0^{-1}}|^{TS} + \frac{\lambda}{2}\frac{\tau_0^2}{(\tau_{AF} + \tau_0)}$$

and the value of the partial derivative for firms with high fundamental uncertainty, or $\tau_0^{-1} = \tau_{0,h}^{-1}$ is therefore

$$\frac{\partial C^*}{\partial \tau_0^{-1}}|_{\tau_{0,h}^{-1}}^{TS} = \frac{1}{2}\left[ln\left(\frac{r}{\lambda_h}\right) - ln\left(\tau_{AF} + \tau_{0,h}\right) - 1\right]\frac{\partial \lambda}{\partial \tau_0^{-1}}|_{\tau_{0,h}^{-1}}^{TS} + \frac{\lambda_h}{2}\frac{1}{\left(\tau_{AF}\tau_{0,h}^{-2} + \tau_{0,h}^{-1}\right)}$$

and for firms with low fundamental uncertainty, or $\tau_0^{-1} = \tau_{0,l}^{-1}$, we have

$$\frac{\partial C^*}{\partial \tau_0^{-1}}|_{\tau_{0,l}^{-1}}^{TS} = \frac{1}{2}\left[ln\left(\frac{r}{\lambda_l}\right) - ln\left(\tau_{AF} + \tau_{0,l}\right) - 1\right]\frac{\partial \lambda}{\partial \tau_0^{-1}}|_{\tau_{0,l}^{-1}}^{TS} + \frac{\lambda_l}{2}\frac{1}{\left(\tau_{AF}\tau_{0,l}^{-2} + \tau_{0,l}^{-1}\right)}$$

when the information cost channel is weak for temporary time-series spikes, or

$$\frac{\partial \lambda}{\partial \tau_0^{-1}}|_{\tau_{0,h}^{-1}}^{TS} \approx \frac{\partial \lambda}{\partial \tau_0^{-1}}|_{\tau_{0,l}^{-1}}^{TS}$$

and the cross-sectional effects of the information cost channel is strong, i.e. $\frac{\partial \lambda}{\partial \tau_0^{-1}}|_{\tau_{0,h}^{-1}}^{CS} >> \frac{\partial \lambda}{\partial \tau_0^{-1}}|_{\tau_{0,l}^{-1}}^{CS}$, so that

$$\lambda_h >> \lambda_l$$

we have the attention increase is bigger for firms with higher fundamental uncertainty than for firms with lower fundamental uncertainty, or

$$\frac{\partial C^*}{\partial \tau_0^{-1}}|_{\tau_{0,h}^{-1}}^{TS} > \frac{\partial C^*}{\partial \tau_0^{-1}}|_{\tau_{0,l}^{-1}}^{TS}$$

## B.4 Robustness Figures and Tables

Table B.4: Information Demand: Controlling for Size and Firm Age

This table runs a probit regression to analyze the relationship between AIA and firm uncertainty. We use a probit model to regress the daily AIA on OIV, Abnormal OIV, HLtH, or Abnormal HLtH. OIV and Abnormal OIV is winsorized at the 1 and 99th percentile. HLtH and Abnormal HLtH is winsorized at the 99th percentile. For days when Bloomberg has no value, we do not fill the missing information of AIA with zero. Each regression includes a control for Size and -LN(Firm Age). Standard errors are clustered by firm and day. Statistical significance is denoted as ***, **, and * for $p<0.10$, $p<0.05$, and $p<0.01$, respectively. The sample period is from March 2010 to December 2019.

| | (1) | (2) | (3) | (4) |
|---|---|---|---|---|
| OIV | 1.207*** | | | |
| | (30.2) | | | |
| Ab. OIV | | 0.550*** | | |
| | | (21.8) | | |
| HLtH | | | 25.764*** | |
| | | | (53.1) | |
| Ab. HLtH | | | | 0.921*** |
| | | | | (59.0) |
| Size | 0.236*** | 0.181*** | 0.280*** | 0.202*** |
| | (45.3) | (32.7) | (56.7) | (33.6) |
| -LN(Age) | 0.001 | 0.037*** | -0.023*** | 0.038*** |
| | (0.2) | (5.4) | (-3.7) | (5.1) |
| Observations | 2647050 | 2646269 | 2681426 | 2681426 |
| Pseudo $R^2$ | 0.042 | 0.035 | 0.111 | 0.135 |

Table B.5: Uncertainty and the Return Predictability of Analysts' Conditional Biases

These tables present the Fama French Five Factor alphas of the EHB Q1-Q5 portfolios by uncertainty terciles using our six monthly uncertainty measures. The EHB Q1-Q5 portfolios based on OIV , IVOL, IVOL$_{MA36}$, and abnormal IVOL, are made by cross-sectionally sorting companies first into deciles based on their market value of equity, then conditionally cross-sectionally sorting firms into terciles based on each uncertainty measure. Then firms are then conditionally cross-sectionally sorted into quintiles based on EHB. The EHB Q1-Q5 portfolios based on VIX, and EPU are made by sorting observations into terciles in the time series based on each uncertainty measure. Then, within each cross section, we first divide firms into the 3 size groups from Table 3.2. We finally conditionally cross-sectionally sort firms into quintiles based on EHB. Within each cross section, we calculate the value weighted return within each size-EHB group, and then average across the three size groups. Portfolios are value weighted and are re-balanced on a monthly basis. Q1 (Q5) contains firms with the lowest (highest) values of EHB. T1 (T3) of each uncertainty tercile contains firms with the lowest (highest) values. Panels A through F show the results IVOL, OIV, IVOL$_{MA36}$, Abnormal IVOL, VIX, and EPU, respectively. Standard errors of the resulting regression coefficients are computed based on Newey and West, 1987 with 12 lags. Statistical significance is denoted as ***, **, and * for $p<0.10$, $p<0.05$, and $p<0.01$, respectively. The sample period is from June 1990 to December 2019, with the exception of OIV which begins in January 1996.

Panel A: IVOL

| | EHB Q1 | EHB Q2 | EHB Q3 | EHB Q4 | EHB Q5 | EHB Q1-Q5 |
|---|---|---|---|---|---|---|
| IVOL T1 | 0.031 | 0.114* | -0.000 | 0.045 | 0.032 | -0.000 |
| | (0.39) | (1.75) | (-0.00) | (0.66) | (0.37) | (-0.00) |
| IVOL T2 | 0.207** | 0.017 | -0.124* | -0.124 | -0.308** | 0.514*** |
| | (2.22) | (0.22) | (-1.87) | (-1.41) | (-2.39) | (2.65) |
| IVOL T3 | 0.552*** | 0.277** | -0.162* | -0.150 | -0.534*** | 1.086*** |
| | (2.85) | (2.13) | (-1.74) | (-1.17) | (-2.85) | (3.34) |
| IVOL T1-T3 | -0.520** | -0.163 | 0.162 | 0.196 | 0.566*** | -1.086*** |
| | (-2.34) | (-0.95) | (1.31) | (1.34) | (3.32) | (-3.60) |

Panel B: OIV

| | EHB Q1 | EHB Q2 | EHB Q3 | EHB Q4 | EHB Q5 | EHB Q1-Q5 |
|---|---|---|---|---|---|---|
| OIV T1 | 0.033 | 0.100 | -0.060 | -0.001 | -0.096 | 0.129 |
| | (0.34) | (1.41) | (-0.75) | (-0.01) | (-0.79) | (0.93) |
| OIV T2 | 0.198* | 0.087 | -0.088 | -0.120 | -0.294** | 0.492** |
| | (1.69) | (1.41) | (-1.03) | (-1.22) | (-2.07) | (2.37) |
| OIV T3 | 0.364* | 0.291** | -0.029 | -0.141 | -0.681*** | 1.045** |
| | (1.71) | (2.27) | (-0.25) | (-0.84) | (-2.79) | (2.57) |
| OIV T1-T3 | -0.331 | -0.190 | -0.031 | 0.140 | 0.585** | -0.916** |
| | (-1.30) | (-1.09) | (-0.22) | (0.84) | (2.55) | (-2.49) |

Panel C: $IVOL_{MA36}$

| | EHB Q1 | EHB Q2 | EHB Q3 | EHB Q4 | EHB Q5 | EHB Q1-Q5 |
|---|---|---|---|---|---|---|
| $\text{IVOL}_{MA36}$ T1 | -0.003 | 0.061 | -0.040 | 0.022 | -0.036 | 0.033 |
| | (-0.03) | (0.96) | (-0.55) | (0.31) | (-0.31) | (0.22) |
| $\text{IVOL}_{MA36}$ T2 | 0.102 | -0.051 | -0.061 | -0.142* | -0.235 | 0.338* |
| | (1.33) | (-0.63) | (-0.96) | (-1.66) | (-1.48) | (1.79) |
| $\text{IVOL}_{MA36}$ T3 | 0.430** | 0.398*** | -0.052 | -0.130 | -0.315 | 0.746** |
| | (2.21) | (2.74) | (-0.60) | (-1.05) | (-1.49) | (2.30) |
| $\text{IVOL}_{MA36}$ T1-T3 | -0.433* | -0.336* | 0.011 | 0.152 | 0.279 | -0.712** |
| | (-1.73) | (-1.85) | (0.09) | (1.04) | (1.26) | (-2.13) |

Panel D: Abnormal IVOL

| | EHB Q1 | EHB Q2 | EHB Q3 | EHB Q4 | EHB Q5 | EHB Q1-Q5 |
|---|---|---|---|---|---|---|
| Ab. IVOL T1 | 0.382*** | 0.255*** | 0.065 | 0.014 | -0.274** | 0.656*** |
| | (3.82) | (3.57) | (1.05) | (0.19) | (-1.97) | (3.44) |
| Ab. IVOL T2 | 0.136 | 0.068 | 0.014 | -0.009 | -0.224 | 0.360 |
| | (1.28) | (0.97) | (0.23) | (-0.10) | (-1.43) | (1.62) |
| Ab. IVOL T3 | 0.010 | -0.004 | -0.167* | -0.065 | -0.374* | 0.384 |
| | (0.11) | (-0.05) | (-1.85) | (-0.59) | (-1.77) | (1.60) |
| Ab. IVOL T1-T3 | 0.373*** | 0.259** | 0.233* | 0.080 | 0.100 | 0.273 |
| | (2.73) | (2.35) | (1.89) | (0.53) | (0.49) | (1.37) |

Panel E: VIX

| | EHB Q1 | EHB Q2 | EHB Q3 | EHB Q4 | EHB Q5 | EHB Q1-Q5 |
|---|---|---|---|---|---|---|
| VIX T1 | 0.210** | 0.141** | -0.061 | -0.033 | -0.293** | 0.503** |
| | (2.39) | (2.19) | (-1.25) | (-0.69) | (-2.17) | (2.57) |
| VIX T2 | 0.306*** | 0.182** | -0.014 | -0.152* | -0.470*** | 0.777*** |
| | (5.08) | (2.28) | (-0.27) | (-1.81) | (-3.43) | (4.40) |
| VIX T3 | 0.209 | 0.047 | -0.014 | 0.050 | 0.069 | 0.140 |
| | (1.15) | (0.50) | (-0.18) | (0.41) | (0.24) | (0.34) |
| VIX T1-T3 | 0.000 | 0.094 | -0.047 | -0.083 | -0.362 | 0.363 |
| | (0.00) | (0.82) | (-0.49) | (-0.63) | (-1.13) | (0.80) |

Panel F:EPU

| | EHB Q1 | EHB Q2 | EHB Q3 | EHB Q4 | EHB Q5 | EHB Q1-Q5 |
|---|---|---|---|---|---|---|
| EPU T1 | 0.536*** | 0.156** | 0.002 | -0.294*** | -0.708*** | 1.244*** |
| | (3.59) | (2.32) | (0.05) | (-4.20) | (-3.64) | (4.09) |
| EPU T2 | 0.258*** | 0.192*** | -0.121 | -0.135 | -0.470*** | 0.728*** |
| | (3.39) | (3.33) | (-1.48) | (-1.23) | (-2.66) | (3.10) |
| EPU T3 | 0.169 | 0.071 | 0.018 | 0.118 | -0.061 | 0.230 |
| | (1.57) | (1.08) | (0.22) | (1.00) | (-0.32) | (0.88) |
| EPU T1-T3 | 0.367** | 0.085 | -0.016 | -0.412*** | -0.647** | 1.014** |
| | (1.97) | (0.94) | (-0.17) | (-3.01) | (-2.39) | (2.52) |

Table B.6: Uncertainty and the Return Predictability of Analysts' Conditional Biases

This table presents the results of Fama-MacBeth monthly regressions of one-month-ahead returns (in percentage points) on the normalized rank of EHB (i.e., the rank scaled by the number of stocks in the cross-section). Standard errors of the resulting Fama-MacBeth regression coefficient are computed based on Newey and West, 1987 with 12 lags. Statistical significance is denoted as ***, **, and * for p<0.10, p<0.05, and p<0.01, respectively. Columns (1) to (4) in Panels A to C show the regression results based on cross-sectional terciles of OIV , IVOL, and $\text{IVOL}_{MA36}$ orthogonalized to size, respectively. Terciles are generated by first cross-sectionally regressing the normalized rank of the uncertainty measure on the normalized rank of size. Then firms are cross-sectionally sorted into terciles from the residual of the regression. Columns (5) to (8) in Panels A to C show the regression results in subsamples based on the time-series terciles of VIX and EPU as well as the cross-sectional terciles of Abnormal IVOL, respectively. All regressions include a the normalized rank of Size as a control. T1 indicates the lowest values and T3 the highest values. The sample period is from June 1990 to December 2019, with the exception of OIV which begins in January 1996.

Panel A: OIV and VIX

| | (1) OIV T1 | (2) OIV T2 | (3) OIV T3 | (4) T3-T1 | (5) VIX T1 | (6) VIX T2 | (7) VIX T3 | (8) T3-T1 |
|---|---|---|---|---|---|---|---|---|
| EHB | 0.012 | -0.273 | -1.246** | -1.258*** | -0.556** | -1.145** | -0.123 | 0.433 |
| | (0.1) | (-1.1) | (-2.5) | (-2.9) | (-2.2) | (-2.5) | (-0.2) | (0.7) |
| Observations | 151649 | 151957 | 151678 | 303327 | 212688 | 211438 | 211102 | 423790 |

Panel B: IVOL and EPU

| | (1) IVOL T1 | (2) IVOL T2 | (3) IVOL T3 | (4) T3-T1 | (5) EPU T1 | (6) EPU T2 | (7) EPU T3 | (8) T3-T1 |
|---|---|---|---|---|---|---|---|---|
| EHB | 0.060 | -0.321 | -1.299*** | -1.358*** | -1.174*** | -1.032** | 0.492 | 1.667** |
| | (0.4) | (-1.5) | (-3.4) | (-3.8) | (-3.3) | (-2.2) | (0.9) | (2.3) |
| Observations | 204665 | 211841 | 212264 | 416929 | 236916 | 214029 | 184283 | 421199 |

Panel C: $\text{IVOL}_{MA36}$ and Abnormal IVOL

| | (1) $\text{IVOL}_{MA36}$ T1 | (2) $\text{IVOL}_{MA36}$ T2 | (3) $\text{IVOL}_{MA36}$ T3 | (4) T3-T1 | (5) Ab. IVOL T1 | (6) Ab. IVOL T2 | (7) Ab. IVOL T3 | (8) T3-T1 |
|---|---|---|---|---|---|---|---|---|
| EHB | 0.071 | -0.083 | -1.016*** | -1.087*** | -0.697*** | -0.243 | -0.264 | 0.432* |
| | (0.3) | (-0.4) | (-3.1) | (-3.8) | (-3.2) | (-1.1) | (-0.9) | (1.9) |
| Observations | 184755 | 191599 | 192297 | 377052 | 189273 | 190603 | 188775 | 378048 |

Table B.7: Information Demand Terciles: Natural Log

This table runs a probit regression to analyze the relationship between AIA and firm uncertainty. We use a probit model to regress the daily AIA on LN(1+Abnormal OIV) or LN(1+Abnormal HLtH) for firms sorted into long-run uncertainty terciles. The terciles are generated by orthogonalizing the cross-sectional normalized rank of $IVOL_{MA36}$ by to the cross-sectional normalized rank of Size. This is done on a monthly basis and firms are sorted into terciles based on the values generated at the end of the previous month. Abnormal OIV is winsorized at the 1 and 99th percentile. Abnormal HLtH is winsorized at the 99th percentile. For days when Bloomberg has no value, we do not fill the missing information of AIA with zero. Each regression includes a control for Size. Standard errors are clustered by firm and day. Statistical significance is denoted as ***, **, and * for $p<0.10$, $p<0.05$, and $p<0.01$, respectively. The sample period is from March 2010 to December 2019.

| | (1)<br>T1 | (2)<br>T2 | (3)<br>T3 | (4)<br>T1-T3 | (5)<br>T1 | (6)<br>T2 | (7)<br>T3 | (8)<br>T1-T3 |
|---|---|---|---|---|---|---|---|---|
| LN(1+Ab. OIV) | 0.868*** | 1.164*** | 1.271*** | -0.403*** | | | | |
| | (13.8) | (18.3) | (20.8) | (-6.1) | | | | |
| LN(1+Ab. HLtH) | | | | | 1.801*** | 2.114*** | 2.392*** | -0.591*** |
| | | | | | (51.8) | (64.0) | (71.8) | (-18.9) |
| Size | 0.199*** | 0.193*** | 0.174*** | 0.025** | 0.214*** | 0.215*** | 0.199*** | 0.016 |
| | (28.4) | (29.7) | (23.3) | (2.6) | (28.9) | (30.4) | (23.9) | (1.5) |
| Observations | 803138 | 878222 | 964909 | 1768047 | 819325 | 889614 | 972487 | 1791812 |
| Pseudo $R^2$ | 0.049 | 0.042 | 0.028 | 0.039 | 0.120 | 0.140 | 0.147 | 0.137 |

Table B.8: Uncertainty and the Return Predictability of Analysts' Conditional Biases

This figure shows the Fama French Five Factor alphas of the EHB Q1-Q5 portfolios by uncertainty terciles using our six monthly uncertainty measures. The EHB Q1-Q5 portfolios based on OIV, IVOL, $IVOL_{MA36}$, and Abnormal IVOL, are made by cross-sectionally sorting companies first into deciles based on their market value of equity, then conditionally cross-sectionally sorting firms into terciles based on each uncertainty measure. Then firms are then conditionally cross-sectionally sorted into quintiles based on EHB. The EHB Q1-Q5 portfolios based on VIX, and EPU are made by sorting observations into terciles in the time series based on each uncertainty measure. Then, within each cross section, we first divide firms into the 3 size groups from Table 3.2. We finally conditionally cross-sectionally sort firms into quintiles based on EHB. Within each cross section, we calculate the value weighted return within each size-EHB group, and then average across the three size groups. Portfolios are value weighted and are re-balanced on a monthly basis. Q1 (Q5) contains firms with the lowest (highest) values of EHB. T1 (T3) of each uncertainty tercile contains firms with the lowest (highest) values. Standard errors of the resulting regression coefficients are computed based on Newey and West, 1987 with 12 lags. Statistical significance is denoted as ***, **, and * for $p<0.10$, $p<0.05$, and $p<0.01$, respectively. Q1 indicates the lowest values and Q5 the highest values for EHB. The sample period is from June 1990 to December 2019, with the exception of OIV which begins in January 1996.

| | OIV | IVOL | $IVOL_{MA36}$ | Ab. IVOL | VIX | EPU |
|---|---|---|---|---|---|---|
| T1 | 0.129 | -0.000 | 0.033 | 0.656*** | 0.503** | 1.244*** |
| | (0.93) | (-0.00) | (0.22) | (3.44) | (2.57) | (4.09) |
| T2 | 0.492** | 0.514*** | 0.338* | 0.360 | 0.777*** | 0.728*** |
| | (2.37) | (2.65) | (1.79) | (1.62) | (4.40) | (3.10) |
| T3 | 1.045** | 1.086*** | 0.746** | 0.384 | 0.140 | 0.230 |
| | (2.57) | (3.34) | (2.30) | (1.60) | (0.34) | (0.88) |
| T1-T3 | -0.916** | -1.086*** | -0.712** | 0.273 | 0.363 | 1.014** |
| | (-2.49) | (-3.60) | (-2.13) | (1.37) | (0.80) | (2.52) |

Table B.9: OIV and the Return Predictability of Analysts' Conditional Biases

This table presents the excess returns and Fama French Five Factor alphas for double sort portfolios created by cross-sectionally sorting companies into terciles based on OIV then conditionally cross-sectionally sorting firms into quintiles based on their EHB. We first divide firms into the 3 size groups from table 3.2. Then, within each group, we cross-sectionally sort companies into terciles based on OIV, before finally conditionally sorting firms into quintiles based on their EHB. Then, we calculate the value weighted average return within each size-OIV-EHB group by month, before averaging across the 3 size groups within each OIV-EHB portfolio. Panel A shows the excess return results and panel B shows the FF5 results. Portfolios are value weighted and are re-balanced on a monthly basis. Standard errors of the resulting regression coefficients are computed based on Newey and West, 1987 with 12 lags. Statistical significance is denoted as ***, **, and * for $p<0.10$, $p<0.05$, and $p<0.01$, respectively. T1 indicates the lowest values and T3 the highest values for OIV. Q1 indicates the lowest values and Q5 the highest values for EHB. The sample period is from January 1996 to December 2019.

Panel A: Excess Return

| | EHB Q1 | EHB Q2 | EHB Q3 | EHB Q4 | EHB Q5 | EHB Q1-Q5 |
|---|---|---|---|---|---|---|
| OIV T1 | 0.765*** | 0.807*** | 0.679*** | 0.789*** | 0.768*** | -0.003 |
| | (3.40) | (3.92) | (3.21) | (3.65) | (2.96) | (-0.02) |
| OIV T2 | 0.949*** | 0.821*** | 0.762*** | 0.734*** | 0.660* | 0.290 |
| | (3.03) | (2.88) | (2.67) | (2.61) | (1.92) | (1.36) |
| OIV T3 | 0.934* | 0.832* | 0.562 | 0.541 | -0.005 | 0.939** |
| | (1.80) | (1.65) | (1.20) | (1.06) | (-0.01) | (2.32) |
| OIV T1-T3 | -0.170 | -0.024 | 0.116 | 0.249 | 0.773 | -0.942** |
| | (-0.35) | (-0.05) | (0.30) | (0.57) | (1.62) | (-2.53) |

Panel B: Abnormal Return

| | EHB Q1 | EHB Q2 | EHB Q3 | EHB Q4 | EHB Q5 | EHB Q1-Q5 |
|---|---|---|---|---|---|---|
| OIV T1 | 0.033 | 0.100 | -0.060 | -0.001 | -0.096 | 0.129 |
| | (0.34) | (1.41) | (-0.75) | (-0.01) | (-0.79) | (0.93) |
| OIV T2 | 0.198* | 0.087 | -0.088 | -0.120 | -0.294** | 0.492** |
| | (1.69) | (1.41) | (-1.03) | (-1.22) | (-2.07) | (2.37) |
| OIV T3 | 0.364* | 0.291** | -0.029 | -0.141 | -0.681*** | 1.045** |
| | (1.71) | (2.27) | (-0.25) | (-0.84) | (-2.79) | (2.57) |
| OIV T1-T3 | -0.331 | -0.190 | -0.031 | 0.140 | 0.585** | -0.916** |
| | (-1.30) | (-1.09) | (-0.22) | (0.84) | (2.55) | (-2.49) |

Table B.10: IVOL and the Return Predictability of Analysts' Conditional Biases

This table presents the excess returns and Fama French Five Factor alphas for double sort portfolios created by cross-sectionally sorting companies into terciles based on IVOL then conditionally cross-sectionally sorting firms into quintiles based on their EHB. We first divide firms into the 3 size groups from table 3.2. Then, within each group, we cross-sectionally sort companies into terciles based on IVOL, before finally conditionally sorting firms into quintiles based on their EHB. Then, we calculate the value weighted average return within each size-IVOL-EHB group by month, before averaging across the 3 size groups within each IVOL-EHB portfolio. Panel A shows the excess return results and panel B shows the FF5 results. Portfolios are value weighted and are re-balanced on a monthly basis. Standard errors of the resulting regression coefficients are computed based on Newey and West, 1987 with 12 lags. Statistical significance is denoted as ***, **, and * for p<0.10, p<0.05, and p<0.01, respectively. T1 indicates the lowest values and T3 the highest values for IVOL. Q1 indicates the lowest values and Q5 the highest values for EHB. The sample period is from June 1990 to December 2019.

Panel A: Excess Return

| | EHB Q1 | EHB Q2 | EHB Q3 | EHB Q4 | EHB Q5 | EHB Q1-Q5 |
|---|---|---|---|---|---|---|
| IVOL T1 | 0.805*** | 0.883*** | 0.781*** | 0.879*** | 0.955*** | -0.150 |
| | (3.98) | (4.91) | (4.33) | (4.58) | (4.08) | (-1.14) |
| IVOL T2 | 1.006*** | 0.813*** | 0.776*** | 0.801*** | 0.686** | 0.320* |
| | (4.31) | (3.58) | (3.56) | (3.30) | (2.36) | (1.73) |
| IVOL T3 | 1.100*** | 0.781** | 0.399 | 0.460 | 0.223 | 0.877*** |
| | (2.59) | (2.01) | (1.02) | (1.14) | (0.44) | (2.84) |
| IVOL T1-T3 | -0.295 | 0.101 | 0.382 | 0.419 | 0.732* | -1.028*** |
| | (-0.80) | (0.30) | (1.20) | (1.26) | (1.88) | (-3.63) |

Panel B: Abnormal Return

| | EHB Q1 | EHB Q2 | EHB Q3 | EHB Q4 | EHB Q5 | EHB Q1-Q5 |
|---|---|---|---|---|---|---|
| IVOL T1 | 0.031 | 0.114* | -0.000 | 0.045 | 0.032 | -0.000 |
| | (0.39) | (1.75) | (-0.00) | (0.66) | (0.37) | (-0.00) |
| IVOL T2 | 0.207** | 0.017 | -0.124* | -0.124 | -0.308** | 0.514*** |
| | (2.22) | (0.22) | (-1.87) | (-1.41) | (-2.39) | (2.65) |
| IVOL T3 | 0.552*** | 0.277** | -0.162* | -0.150 | -0.534*** | 1.086*** |
| | (2.85) | (2.13) | (-1.74) | (-1.17) | (-2.85) | (3.34) |
| IVOL T1-T3 | -0.520** | -0.163 | 0.162 | 0.196 | 0.566*** | -1.086*** |
| | (-2.34) | (-0.95) | (1.31) | (1.34) | (3.32) | (-3.60) |

Table B.11: $\text{IVOL}_{MA36}$ and the Return Predictability of Analysts' Conditional Biases

This table presents the excess returns and Fama French Five Factor alphas for double sort portfolios created by cross-sectionally sorting companies into terciles based on $\text{IVOL}_{MA36}$ then conditionally cross-sectionally sorting firms into quintiles based on their EHB. We first divide firms into the 3 size groups from table 3.2. Then, within each group, we cross-sectionally sort companies into terciles based on $\text{IVOL}_{MA36}$ , before finally conditionally sorting firms into quintiles based on their EHB. Then, we calculate the value weighted average return within each size-I$\text{IVOL}_{MA36}$ -EHB group by month, before averaging across the 3 size groups within each $\text{IVOL}_{MA36}$ -EHB portfolio. Panel A shows the excess return results and panel B shows the FF5 results. Portfolios are value weighted and are re-balanced on a monthly basis. Standard errors of the resulting regression coefficients are computed based on Newey and West, 1987 with 12 lags. Statistical significance is denoted as ***, **, and * for $p<0.10$, $p<0.05$, and $p<0.01$, respectively. T1 indicates the lowest values and T3 the highest values for $\text{IVOL}_{MA36}$ . Q1 indicates the lowest values and Q5 the highest values for EHB. The sample period is from June 1990 to December 2019.

Panel A: Excess Return

| | EHB Q1 | EHB Q2 | EHB Q3 | EHB Q4 | EHB Q5 | EHB Q1-Q5 |
|---|---|---|---|---|---|---|
| $\text{IVOL}_{MA36}$ T1 | 0.757*** | 0.826*** | 0.759*** | 0.872*** | 0.915*** | -0.158 |
| | (3.72) | (4.58) | (4.13) | (4.31) | (3.30) | (-0.95) |
| $\text{IVOL}_{MA36}$ T2 | 0.928*** | 0.773*** | 0.835*** | 0.802*** | 0.851*** | 0.076 |
| | (4.07) | (3.53) | (3.86) | (3.36) | (2.60) | (0.37) |
| $\text{IVOL}_{MA36}$ T3 | 1.095*** | 0.962*** | 0.602* | 0.630* | 0.497 | 0.597** |
| | (2.93) | (2.70) | (1.80) | (1.74) | (1.21) | (2.31) |
| $\text{IVOL}_{MA36}$ T1-T3 | -0.338 | -0.136 | 0.157 | 0.241 | 0.418 | -0.756*** |
| | (-0.98) | (-0.44) | (0.56) | (0.86) | (1.43) | (-3.15) |

Panel B: Abnormal Return

| | EHB Q1 | EHB Q2 | EHB Q3 | EHB Q4 | EHB Q5 | EHB Q1-Q5 |
|---|---|---|---|---|---|---|
| $\text{IVOL}_{MA36}$ T1 | -0.003 | 0.061 | -0.040 | 0.022 | -0.036 | 0.033 |
| | (-0.03) | (0.96) | (-0.55) | (0.31) | (-0.31) | (0.22) |
| $\text{IVOL}_{MA36}$ T2 | 0.102 | -0.051 | -0.061 | -0.142* | -0.235 | 0.338* |
| | (1.33) | (-0.63) | (-0.96) | (-1.66) | (-1.48) | (1.79) |
| $\text{IVOL}_{MA36}$ T3 | 0.430** | 0.398*** | -0.052 | -0.130 | -0.315 | 0.746** |
| | (2.21) | (2.74) | (-0.60) | (-1.05) | (-1.49) | (2.30) |
| $\text{IVOL}_{MA36}$ T1-T3 | -0.433* | -0.336* | 0.011 | 0.152 | 0.279 | -0.712** |
| | (-1.73) | (-1.85) | (0.09) | (1.04) | (1.26) | (-2.13) |

Table B.12: Abnormal IVOL and the Return Predictability of Analysts' Conditional Biases

This table presents the excess returns and Fama French Five Factor alphas for double sort portfolios created by cross-sectionally sorting companies into terciles based on Abnormal IVOL then conditionally cross-sectionally sorting firms into quintiles based on their EHB. We first divide firms into the 3 size groups from table 3.2. Then, within each group, we cross-sectionally sort companies into terciles based on Abnormal IVOL, before finally conditionally sorting firms into quintiles based on their EHB. Then, we calculate the value weighted average return within each size-Abnormal IVOL-EHB group by month, before averaging across the 3 size groups within each Abnormal IVOL-EHB portfolio. Panel A shows the excess return results and panel B shows the FF5 results. Portfolios are value weighted and are re-balanced on a monthly basis. Standard errors of the resulting regression coefficients are computed based on Newey and West, 1987 with 12 lags. Statistical significance is denoted as ***, **, and * for $p<0.10$, $p<0.05$, and $p<0.01$, respectively. T1 indicates the lowest values and T3 the highest values for Abnormal IVOL. Q1 indicates the lowest values and Q5 the highest values for EHB. The sample period is from June 1990 to December 2019.

Panel A: Excess Return

| | EHB Q1 | EHB Q2 | EHB Q3 | EHB Q4 | EHB Q5 | EHB Q1-Q5 |
|---|---|---|---|---|---|---|
| Abnormal IVOL T1 | 1.120*** | 0.954*** | 0.795*** | 0.830*** | 0.617** | 0.503*** |
| | (4.90) | (4.50) | (3.78) | (3.38) | (2.04) | (2.90) |
| Abnormal IVOL T2 | 0.889*** | 0.844*** | 0.827*** | 0.855*** | 0.710** | 0.179 |
| | (3.57) | (4.12) | (3.95) | (3.81) | (2.45) | (0.94) |
| Abnormal IVOL T3 | 0.753*** | 0.722*** | 0.659*** | 0.813*** | 0.624 | 0.128 |
| | (3.16) | (3.29) | (2.70) | (2.90) | (1.53) | (0.51) |
| Abnormal IVOL T1-T3 | 0.367*** | 0.231** | 0.135 | 0.016 | -0.007 | 0.374** |
| | (2.77) | (2.41) | (1.16) | (0.12) | (-0.03) | (1.96) |

Panel B: Abnormal Return

| | EHB Q1 | EHB Q2 | EHB Q3 | EHB Q4 | EHB Q5 | EHB Q1-Q5 |
|---|---|---|---|---|---|---|
| Abnormal IVOL T1 | 0.382*** | 0.255*** | 0.065 | 0.014 | -0.274** | 0.656*** |
| | (3.82) | (3.57) | (1.05) | (0.19) | (-1.97) | (3.44) |
| Abnormal IVOL T2 | 0.136 | 0.068 | 0.014 | -0.009 | -0.224 | 0.360 |
| | (1.28) | (0.97) | (0.23) | (-0.10) | (-1.43) | (1.62) |
| Abnormal IVOL T3 | 0.010 | -0.004 | -0.167* | -0.065 | -0.374* | 0.384 |
| | (0.11) | (-0.05) | (-1.85) | (-0.59) | (-1.77) | (1.60) |
| Abnormal IVOL T1-T3 | 0.373*** | 0.259** | 0.233* | 0.080 | 0.100 | 0.273 |
| | (2.73) | (2.35) | (1.89) | (0.53) | (0.49) | (1.37) |

Table B.13: VIX and the Return Predictability of Analysts' Conditional Biases

This table presents the excess returns and Fama French Five Factor alphas for double sort portfolios created by cross-sectionally sorting companies into terciles based on VIX then conditionally cross-sectionally sorting firms into quintiles based on their EHB. We first sort firms into terciles based on VIX. Then with each tercile, we cross-sectionally divide firms into 3 size groups from table 3.2. Then firms are finally conditionally sorted into quintiles based on their EHB. Then, we calculate the value weighted average return within each size-VIX-EHB group by month, before averaging across the 3 size groups within each VIX-EHB portfolio. Panel A shows the excess return results and panel B shows the FF5 results. Portfolios are value weighted and are re-balanced on a monthly basis. Standard errors of the resulting regression coefficients are computed based on Newey and West, 1987 with 12 lags. Statistical significance is denoted as ***, **, and * for $p<0.10$, $p<0.05$, and $p<0.01$, respectively. T1 indicates the lowest values and T3 the highest values for VIX. Q1 indicates the lowest values and Q5 the highest values for EHB. The sample period is from June 1990 to December 2019.

Panel A: Excess Return

| | EHB Q1 | EHB Q2 | EHB Q3 | EHB Q4 | EHB Q5 | EHB Q1-Q5 |
|---|---|---|---|---|---|---|
| VIX T1 | 0.652*** | 0.546*** | 0.392** | 0.453** | 0.276 | 0.376* |
| | (3.23) | (2.70) | (1.98) | (2.08) | (0.95) | (1.84) |
| VIX T2 | 1.476*** | 1.184*** | 1.097*** | 0.963*** | 0.702 | 0.774*** |
| | (5.41) | (4.26) | (4.09) | (2.70) | (1.46) | (2.62) |
| VIX T3 | 0.821* | 0.692 | 0.652 | 0.785* | 0.740 | 0.081 |
| | (1.74) | (1.61) | (1.51) | (1.65) | (1.18) | (0.20) |
| VIX T1-T3 | 0.000 | 0.094 | -0.047 | -0.083 | -0.362 | 0.363 |
| | (0.00) | (0.82) | (-0.49) | (-0.63) | (-1.13) | (0.80) |

Panel B: Abnormal Return

| | EHB Q1 | EHB Q2 | EHB Q3 | EHB Q4 | EHB Q5 | EHB Q1-Q5 |
|---|---|---|---|---|---|---|
| VIX T1 | 0.210** | 0.141** | -0.061 | -0.033 | -0.293** | 0.503** |
| | (2.39) | (2.19) | (-1.25) | (-0.69) | (-2.17) | (2.57) |
| VIX T2 | 0.306*** | 0.182** | -0.014 | -0.152* | -0.470*** | 0.777*** |
| | (5.08) | (2.28) | (-0.27) | (-1.81) | (-3.43) | (4.40) |
| VIX T3 | 0.209 | 0.047 | -0.014 | 0.050 | 0.069 | 0.140 |
| | (1.15) | (0.50) | (-0.18) | (0.41) | (0.24) | (0.34) |
| VIX T1-T3 | 0.000 | 0.094 | -0.047 | -0.083 | -0.362 | 0.363 |
| | (0.00) | (0.82) | (-0.49) | (-0.63) | (-1.13) | (0.80) |

Table B.14: EPU and the Return Predictability of Analysts' Conditional Biases

This table presents the excess returns and Fama French Five Factor alphas for double sort portfolios created by cross-sectionally sorting companies into terciles based on EPU then conditionally cross-sectionally sorting firms into quintiles based on their EHB. We first sort firms into terciles based on EPU. Then with each tercile, we cross-sectionally divide firms into 3 size groups from table 3.2. Then firms are finally conditionally sorted into quintiles based on their EHB. Then, we calculate the value weighted average return within each size-EPU-EHB group by month, before averaging across the 3 size groups within each EPU-EHB portfolio. Panel A shows the excess return results and panel B shows the FF5 results. Portfolios are value weighted and are re-balanced on a monthly basis. Standard errors of the resulting regression coefficients are computed based on Newey and West, 1987 with 12 lags. Statistical significance is denoted as ***, **, and * for $p<0.10$, $p<0.05$, and $p<0.01$, respectively. T1 indicates the lowest values and T3 the highest values for EPU. Q1 indicates the lowest values and Q5 the highest values for EHB. The sample period is from June 1990 to December 2019.

Panel A: Excess Return

| | EHB Q1 | EHB Q2 | EHB Q3 | EHB Q4 | EHB Q5 | EHB Q1-Q5 |
|---|---|---|---|---|---|---|
| EPU T1 | 0.869** | 0.498 | 0.464 | 0.301 | 0.080 | 0.789*** |
| | (2.22) | (1.55) | (1.42) | (0.83) | (0.19) | (2.87) |
| EPU T2 | 0.447 | 0.347 | 0.048 | -0.000 | -0.396 | 0.842** |
| | (1.32) | (1.02) | (0.12) | (-0.00) | (-0.59) | (2.26) |
| EPU T3 | 1.631*** | 1.577*** | 1.629*** | 1.902*** | 2.035*** | -0.403 |
| | (3.67) | (3.68) | (3.73) | (3.86) | (3.29) | (-1.09) |
| EPU T1-T3 | 0.367** | 0.085 | -0.016 | -0.412*** | -0.647** | 1.014** |
| | (1.97) | (0.94) | (-0.17) | (-3.01) | (-2.39) | (2.52) |

Panel B: Abnormal Return

| | EHB Q1 | EHB Q2 | EHB Q3 | EHB Q4 | EHB Q5 | EHB Q1-Q5 |
|---|---|---|---|---|---|---|
| EPU T1 | 0.536*** | 0.156** | 0.002 | -0.294*** | -0.708*** | 1.244*** |
| | (3.59) | (2.32) | (0.05) | (-4.20) | (-3.64) | (4.09) |
| EPU T2 | 0.258*** | 0.192*** | -0.121 | -0.135 | -0.470*** | 0.728*** |
| | (3.39) | (3.33) | (-1.48) | (-1.23) | (-2.66) | (3.10) |
| EPU T3 | 0.169 | 0.071 | 0.018 | 0.118 | -0.061 | 0.230 |
| | (1.57) | (1.08) | (0.22) | (1.00) | (-0.32) | (0.88) |
| EPU T1-T3 | 0.367** | 0.085 | -0.016 | -0.412*** | -0.647** | 1.014** |
| | (1.97) | (0.94) | (-0.17) | (-3.01) | (-2.39) | (2.52) |

Table B.15: 1/Age and the Return Predictability of Analysts' Conditional Biases

This table presents the excess returns and Fama French Five Factor alphas for double sort portfolios created by cross-sectionally sorting companies into terciles based on 1/Age then conditionally cross-sectionally sorting firms into quintiles based on their EHB. We first divide firms into the 3 size groups from table 3.2. Then, within each group, we cross-sectionally sort companies into terciles based on 1/Age, before finally conditionally sorting firms into quintiles based on their EHB. Then, we calculate the value weighted average return within each size-1/Age-EHB group by month, before averaging across the 3 size groups within each 1/Age-EHB portfolio. Panel A shows the excess return results and panel B shows the FF5 results. Portfolios are value weighted and are re-balanced on a monthly basis. Standard errors of the resulting regression coefficients are computed based on Newey and West, 1987 with 12 lags. Statistical significance is denoted as ***, **, and * for $p<0.10$, $p<0.05$, and $p<0.01$, respectively. T1 indicates the oldest and T3 the youngest firms. Q1 indicates the lowest values and Q5 the highest values for EHB. The sample period is from June 1990 to December 2019.

Panel A: Excess Return

| | EHB Q1 | EHB Q2 | EHB Q3 | EHB Q4 | EHB Q5 | EHB Q1-Q5 |
|---|---|---|---|---|---|---|
| 1/Age T1 | 0.780*** | 0.743*** | 0.817*** | 0.945*** | 0.851*** | -0.071 |
| | (3.77) | (3.99) | (4.43) | (4.06) | (2.94) | (-0.38) |
| 1/Age T2 | 0.971*** | 0.792*** | 0.813*** | 0.755*** | 0.724** | 0.247 |
| | (4.02) | (3.60) | (3.54) | (3.12) | (2.04) | (1.18) |
| 1/Age T3 | 1.177*** | 0.843*** | 0.614* | 0.477 | 0.173 | 1.004*** |
| | (3.52) | (2.64) | (1.95) | (1.40) | (0.37) | (3.98) |
| 1/Age T1-T3 | -0.397* | -0.100 | 0.204 | 0.468** | 0.678** | -1.075*** |
| | (-1.78) | (-0.43) | (0.89) | (1.99) | (2.06) | (-5.13) |

Panel B: Abnormal Return

| | EHB Q1 | EHB Q2 | EHB Q3 | EHB Q4 | EHB Q5 | EHB Q1-Q5 |
|---|---|---|---|---|---|---|
| 1/Age T1 | -0.019 | -0.047 | -0.017 | 0.020 | -0.246* | 0.227 |
| | (-0.20) | (-0.70) | (-0.31) | (0.26) | (-1.69) | (1.11) |
| 1/Age T2 | 0.227** | 0.074 | 0.058 | -0.038 | -0.203 | 0.430* |
| | (2.42) | (0.99) | (0.85) | (-0.44) | (-1.16) | (1.81) |
| 1/Age T3 | 0.535*** | 0.340*** | 0.061 | -0.127 | -0.500*** | 1.035*** |
| | (3.89) | (4.02) | (0.84) | (-1.19) | (-2.98) | (3.77) |
| 1/Age T1-T3 | -0.554*** | -0.387*** | -0.078 | 0.146 | 0.254** | -0.809*** |
| | (-4.04) | (-3.57) | (-0.80) | (1.18) | (1.96) | (-3.89) |