Essays in Bank Accounting and Regulation

Essays in Bank Accounting and Regulation

A dissertation presented

by

Rajesh Vijayaraghavan

to

The Accounting and Management Department, Harvard Business School

in partial fulfillment of the requirements

for the degree of

Doctor of Business Administration

in the subject of

Accounting and Management

Harvard University

Cambridge, Massachusetts

April 2017

Dissertation Advisors:
Professor Paul Healy
Professor V.G. Narayanan

Author:
Rajesh Vijayaraghavan

Essays in Bank Accounting and Regulation

Abstract

This dissertation comprises of two essays on the accounting rules and regulations. The first essay explores the accounting rules for bank loan loss recognition. Motivated by the FASB's new proposal that introduces an expected loss methodology for recognizing losses, it examines two questions that are related to the current GAAP and the new accounting rule. It develops an empirical model of loan loss prediction from the machine learning literature, and shows that it outperforms current GAAP. It then demonstrates the value of expanding the inputs to the model, as proposed by the new rule. Finally, it examines the drivers of the performance difference between the developed model and the current GAAP. The second essay studies the agency conflict between shareholders and managers of firms. It focuses on the regulation around shareholder activism and the proposals that they submit. In particular, it considers the proposals that managers seek to exclude from the proxy statement. Using a hand-collected data set of SEC "no-action" letters, it documents that the shareholder proposal mechanism has a broader set of components than that considered by prior research, and provides a number of empirical regularities. It further documents that shareholder proposals are part of a larger mosaic of shareholder intervention in companies that often go together.

Contents

Abstract . . . iii
Acknowledgments . . . ix

Introduction **1**

1 Recognizing Loan Losses in Banks: An Examination of Alternative Approaches **3**
1.1 Introduction . . . 3
1.2 Motivation and Background . . . 10
1.2.1 Current GAAP Rules – Implemention . . . 13
1.2.2 CECL Rule and the Use of Broader Information . . . 14
1.2.3 Evaluating the ILM and the CECL . . . 15
1.2.4 Research on Loan Loss Accounting . . . 16
1.3 Empirical Framework . . . 18
1.3.1 Current US GAAP . . . 20
1.3.2 CECL Model . . . 21
1.3.3 Measuring Loan Losses . . . 22
1.3.4 Lasso Approach . . . 23
1.3.5 Predicting Loan Losses . . . 25
1.4 Data . . . 26
1.4.1 Sample selection . . . 26
1.4.2 County Data . . . 27
1.4.3 Bank Data and Summary Statistics . . . 30
1.5 Empirical Results . . . 31
1.5.1 Predicting Allowance from Lasso . . . 33
1.5.2 Evaluation of the Limited Information Model . . . 36
1.5.3 Evaluation of the Full Information Model . . . 37
1.5.4 How much more Allowance? . . . 42
1.5.5 Discussion of the Limited and Full Information Models . . . 44
1.5.6 Allowances in States that Suffered Severe Recession . . . 45
1.5.7 How does the Limited Information Model Perform for Large Banks? . . . 52
1.5.8 How Robust are the Horizon of Future Losses? . . . 56

1.6 What Drives the Performance Difference between LINM and ILM? 56
1.6.1 Examining the Weights on Input Variables by the LINM and ILM 58
1.6.2 Exploiting Variation in Banks Operating in States with Severe and Less Severe Crises 61
1.6.3 The Role of Aggregate Data in the LINM and ILM Models 62
1.6.4 Agency Issues in Recognizing Allowances 66
1.7 Conclusion and Suggestions for Future Research 69

2 What Else do Shareholders Want? Shareholder Proposals Contested by Firm Management 74
2.1 Introduction 74
2.2 Institutional Background 78
2.2.1 Shareholder Proposal Process 78
2.2.2 Lifecycle of a Shareholder Proposal 81
2.2.3 Examples of Contested Shareholder Proposals 83
2.3 Examination of Contested Shareholder Proposals 86
2.3.1 Contested Shareholder Proposal Sample 86
2.3.2 Frequency of Contested Proposals 87
2.3.3 Who Receives and Contests Proposals? 90
2.3.4 Shareholding Levels of Proposal Submitters 95
2.3.5 Contested Proposals that are Withdrawn 100
2.3.6 Contested Proposals that are Placed on the Proxy and Voted 103
2.4 Discussion 109
2.5 Conclusion 111

References 114

Appendix A Appendix to Chapter 1 122
A.1 Current GAAP Rules 122
A.2 Data Sources 122
A.2.1 Variable Definitions 123
A.3 Empirical Model Appendix 123
A.4 Supplementary Tables and Figures 124

Appendix B Appendix to Chapter 2 132
B.1 Timeline for Submitting Proposal 132

List of Tables

1.1 Selection of Sample Banks . . . 28
1.2 Summary Statistics . . . 32
1.3 Variables used in the Models . . . 33
1.4 Allowance Model Cumulative Performance By Year . . . 38
1.5 Allowance Model Performance By Year . . . 39
1.6 Allowance Model Cumulative Performance by Year in US States with Severe Banking Crisis . . . 48
1.7 Allowance Model Performance by Year in US States with Severe Banking Crisis . . . 49
1.8 Coefficients from the LINM and ILM Model of Sample Banks . . . 59

2.1 Shareholder Proposals Received, Contested, and Excluded by Meeting year . . . 88
2.2 Frequency of Reasons for Exclusion of Shareholder Proposals . . . 89
2.3 Types of Shareholder Proposals Received, Contested, and Excluded. . . . 90
2.4 Summary Statistics of Firms Contesting Proposals . . . 91
2.5 Analysis of Firms that Receive and Contest Shareholder Proposals . . . 96
2.6 Share Holdings . . . 100
2.7 Analysis of Contested Proposals that Are Withdrawn. . . . 101
2.8 Likelihood of Any Proposal Receiving Shareholder Approval . . . 105
2.9 Likelihood of Contested Proposals Receiving Shareholder Approval . . . 106
2.10 Level of Shareholder Approval for Contested Proposals . . . 108
2.11 Corporate Changes and Receiving/Contesting Proposals . . . 110

A.1 Data Sources . . . 127
A.2 Variable Definitions . . . 128
A.3 Number of Sample Banks by Counties. . . . 129
A.4 Coefficients from the LINM Lasso Model . . . 130
A.5 Implicit Coefficients from the ILM Model . . . 131

List of Figures

1.1 Summary of Allowance Model Performance. 8
1.2 Ratio of Chargeoff to Allowance 11
1.3 Sample banks by county 29
1.4 Allowance Model Cumulative Performance By Year. 40
1.5 Allowance Model Performance By Year 41
1.6 Difference between Predicted Allowance and Actual allowance 43
1.7 Distribution of Failed Banks in the US. 46
1.8 Allowance Model Cumulative Performance By Year in US States with Severe Banking Crisis 50
1.9 Allowance Model Performance By Year in US States with Severe Banking Crisis . . . 51
1.10 Difference between Predicted Allowance and Actual allowance in US States with Severe Banking Crisis 53
1.11 Summary of Allowance Model Performance for Large Banks 54
1.12 Allowance Model Cumulative Performance By Year for Large Banks 55
1.13 Difference between Predicted Allowance and Actual allowance for Large Banks . . . 57
1.14 LINM vs ILM Decomposed by Banks with Severe Banking Crisis States and Less Severe Banking Crisis States 63
1.15 Allowance Model Cumulative Performance By Year in US States with Less Severe and Severe Banking Crisis 64
1.16 LINM vs ILM for sample county banks operating in states 67
1.17 LINM vs ILM for sample county banks operating in states 68
1.18 Summary of Allowance Model Performance for Large Banks that are Above and Below Capital 70
1.19 Difference between Predicted Allowance and Actual allowance for Large Banks Based on their Degree of Capital Constrain 71

2.1 Shareholder Proposal Process 83
2.2 Time Line for JP Morgan Chase – Risk Management Proposal 84
2.3 Time Line for 3M – Animal Testing Shareholder Proposal 85

A.1 Allowance Model Cumulative Performance By Year. 124

A.2 LINM vs ILM for county banks in BEA Region states 125
A.3 LINM vs ILM for county banks in BEA Region states 126

B.1 Timeline for Shareholder Proposal 133

Acknowledgments

As I walk past the bell at Harvard Business School's Baker lawn every morning to get to my office, there is just one recurring thought. Although the journey thus far has been surreal, the real sense of accomplishment would be felt the moment I cross the bell, and walk up the revered stage in the Baker lawn on commencement day to receive the recognition for the work that has gone into developing this thesis. I have a number of people to thank for guiding and supporting me along the way.

My committee opened the gates for my research agenda. Their dedication towards my research focus is phenomenal. They knew how to challenge me without discouraging me. I am greatly indebted to them. Paul Healy forced me to think more precisely. I would always look forward to the meetings in his office, as I would learn something new every time. He patiently listened to my narratives, pushed me to come up with better ones. He has read more versions of my paper than I have written, and the meticulous feedback he gave me is what every graduate student craves for. The ink that went into writing the feedback has left an indelible influence in me as a scholar and researcher.

V.G. Narayanan set me up on my present research path. He encouraged me to work on banking and is largely responsible for my work in this area. He has been patient, understanding, and a tremendous source of inspiration. He has given me sound advice from scholarly to personal. He demanded research focus and clarity of thinking. VG has been a constant source of support throughout the ups and downs that are part of Doctoral study. He and his wife, Sunanda, have been very kind to my family and me since the day I started the doctoral program.

Ian Gow and I share a mutual interest in computer science, which largely helped in tackling my dissertation question. He invested his time and energy generously in my learning, and helped me frame my work. He is one of the main reasons why the word machine learning (although he likes it to be called statistical learning) shows up in my thesis. He made me an empirical scholar, and instilled in me the notion of what good empirical research is about. The high-energy discussions with him drove me to understand my work far better than I would have on my own.

David Scharfstein has had a major influence on the course of this thesis. I have enjoyed my discussions with him and am amazed by how he gets to the meat of the discussion very early on. He has given support and encouragement right from the first day I met him in his office. It was after my initial discussions with him was I fully convinced that the key elements of my paper were complete. He showed me the art of listening, and has the knack of picking out what is essential. He taught me the value of doing interdisciplinary work. I hope I will have an opportunity to pursue fully many of the ideas we discussed.

In addition, I have drawn inspiration from many faculties whom I have come to know during my years at HBS. They have all been great mentors, as I felt their influence on me. A few people deserve special mention: Srikant Datar, Bob Kaplan, and Krishna Palepu for continuously allowing me to share my research and progress with them. Boris Groysberg for his constant support and care. He is a kind of advisor that many doctoral students will yearn for. Eugene Soltes and Suraj Srinivasan for their contribution to the development of my research craft, and for patiently working with me in correcting my mistakes. Gerardo Perez Cavazos for showing the discipline while building collaborations. Dennis Campbell for many engaging discussions. Bengt Holmstrom for his invaluable guidance on navigating the doctoral program and for sharing his perspective on life. I am an ardent follower of his research and an admirer of his teaching style. I will be ever grateful to Raghu Sundaram and Kamakoti Veezhinathan for their mentorship and encouragement that led me to apply to the doctoral program at HBS.

I have been lucky to have large circle of friends and I am extremely grateful to all of them for being a part of my life. My doctoral student cohort – Carolyn Deller, Akash Chattopadhyay, and Henry Eyring have contributed towards my research and have been a constant source of support. Along with their families, they have been my best friends as well as my confidantes. They have helped me navigate the DBA program whilst swimming in similar waters. Hummy Song, Adi Gupta, JinHyung Kim, and Everett Spain for their precious advise and arming with the arsenal of smiles. Yo–Jud Cheng for many thoughtful discussions. Rohan Kekre and Filippo Mezzanotti have been bounce boards for many of my ideas, and I have enjoyed all my discussions with them ranging from music to food, and football to cricket. Conversations with them led to a number of key insights. I am also thankful to the amazing set of doctoral students I met during my time at the Harvard Business School. Arrvindh Shriraman, a friend from my undergrad days, played a pivotal role in my continued pursuit of a doctoral program. He also played an important role in helping me complete my thesis. Arun and Anusha for being with me in many special moments. Finally, I thank my high school, college, and other friends in New York, Boston, West Coast, and in India for always being there to guide me.

I sincerely thank Jen Mucciarone, John Korn, Marais Young, LuAnn Langan, and Daria Russell who ensured excellent administration of the Doctoral program. They always answered my calls for assistance, and ensured issues were resolved promptly. This dissertation relied heavily on data and programing resources of the Harvard Business School's research computing services. I thank Alex Caracuzzo, Barbara Esty, Andrew Marder, Sarah Woolverton, and James Zeitler for the invaluable suggestions and advice they provided. It has been fun working with all of them.

I am very thankful to the audit firm partners, bank examiners, regulators, and other practitioners for sharing many of their perspectives that helped shape the institutional details in my thesis. It is through the discussions did I further realize the importance of managerial-relevant work as part of research.

My parents have been a constant source of encouragement. They were my first teachers, encouraging my intellectual curiosity, and were supportive of many of my decisions along the way. I thank them immensely for the opportunities they provided me. I also thank my grand parents, uncles, and aunts who all taught me to treasure learning. Heart felt thanks to my brother and his wife for their emotional support. My in laws provided continual encouragement. I am very grateful to them for understanding the trials and tribulations of a graduate student's life and being supportive throughout. My wonderful son Rushil has been my pillar of strength and is my biggest stress buster. He is too young to read these words, but he is the greatest gift I have received. Words are not enough to thank my dear wife, Nirupama. I am proud of her personal accomplishments, having excelled in all her roles. She has been with me in every inch of my travel. She is the lighthouse that kept this ship navigating.

Dedicated to my wonderful family, whose love made it possible to realize my dreams.

Introduction

The dissertation focuses on understanding issues related to two important constituents of modern organizations – the managers and shareholders. The first essay dealing with bank accounting rules for managers investigates the rules for loan loss recognition in banks and ways to improve them. The FASB in June 2016 issued a new rule, effective in December 2019, that will replace current GAAP. The new rule will allow bank managers to use broader information in estimating loan loss allowance. To empirically examine the current GAAP and the new model, the essay exploits the differences in the information sets allowed under the old and the new rules. Using a methodology that combines micro data and machine learning techniques, it provide evidence that it is possible to construct a loan loss recognition model that outperforms the current GAAP without expanding the information set beyond that permitted under the current rule. It finds that expanding this model's information set does not significantly improve its performance. The essay further shows that the difference between the model predicted allowance and the actual allowance recognized by banks is economically meaningful. Using the developed model banks would have recognized larger losses as they entered the financial crisis. The results provide a novel method to evaluate aspects of the new accounting rule before it comes into effect. They suggest that the way information is used, rather than the use of broader information set improves the estimates of loan loss allowance.

The second chapter, joint with Eugene Soltes and Suraj Srinivasan, deals with the regulation around shareholder proposals and their influence on firm strategy and governance. Shareholder proposals provide investors an opportunity to exercise their decision rights within a firm. However, not all proposals created by shareholders receive consideration. Managers have discretion in deciding which proposal to get voted on, and can seek permission from the Securities and Exchange Commission (SEC) to exclude specific proposals from the proxy statement. From 2003–2015, it finds that managers seek to exclude 39% of all proposals they receive, but the SEC does not permit exclusion in over a quarter of the cases. Of the proposals that managers seek to exclude but the SEC does not allow, 21% win shareholder support or the firm voluntarily implements prior to a vote. The analysis of contested shareholder proposals suggests that managers often seek to avoid the implementation of legitimate shareholder interests. The study sheds light on the selection issue on prior studies on shareholder proposal regulation.

Chapter 1

Recognizing Loan Losses in Banks: An Examination of Alternative Approaches

1.1 Introduction

The mortgage crisis has spurred an ongoing debate on the role of transparency in enhancing financial system stability. This has, in turn, revived the longstanding question among policymakers about how banks should account for loan losses on their held for investment loans (see Wall and Koch, 2000; Davenport, 2004). Current US GAAP employs the incurred loss model (ILM), which requires banks to accrue for losses only if there is objective evidence that it is "probable" that a loss event has occurred. The rule applied in practice restricts banks' ability to record losses that are expected, but do not yet meet the probable threshold, allowing use of only a subset of the available information. The FASB issued ASU 2016 – 13, a new accounting standard in June 2016 that replaces the ILM with a rule that (i) eliminates the "probable" threshold, and (ii) broadens the information considered when measuring credit losses to include forward-looking information, which are reasonable and supportable. This rule is referred to as the current expected credit loss methodology (CECL).[1]

In this paper, I empirically examine two questions related to the ILM and CECL models, focusing

[1] See FASB (2016) for further discussion of the new rule.

on the differences in the information sets that they allow. First, I evaluate whether it is possible to construct a predictor of future loan losses that performs better than the ILM approach, without expanding the information set beyond that permitted under that approach. Second, I assess the impact of expanding the information set to include data of the kind proposed under the CECL rule. Together, the empirical analyses contribute to our understanding of the tradeoff between allowing for better use of information and allowing for broader information, and identifies areas where relevant accounting rules can be improved.

Given that allowance estimates directly influence the volatility of bank earnings on its income statements, as well as the value at which loans are reported on the balance sheet, research on loan accounting rules has important implications. Better estimates of loan loss allowances can improve the efficiency of capital allocation. Furthermore, the CECL proposal is widely believed to be the "biggest change ever to bank accounting." [2] The catalyst for the CECL rule was criticism that the ILM delayed recognition of loan losses during the mortgage crisis until they were probable, thereby affecting the adequacy of allowance estimates (Dugan, 2009). In the pre-crisis period, banks had unusually low levels of reserves against eventual loan losses. One possible explanation for this phenomenon is that the ILM limited banks' ability to record adequate provisions for loans that were performing, but that were eventually expected to become delinquent. Under this argument, provisions should have been recorded earlier than allowed by the ILM. However, whether loan loss allowance were untimely because of ILM remains an open empirical question.

Ideally, an evaluation of the performance of the two models would compare the estimated allowance under the CECL model to the allowance under the ILM to identify which is the better predictor of loan losses. While bank financial statements currently report allowances under ILM, there are no data available on the allowance under the CECL. Therefore, my paper uses a prediction model of future net-chargeoffs to obtain counterfactual estimates of allowances and use the root mean squared error to compare their performance.

For my first research question on estimating allowances without expanding the information

[2] See letter from Rob Nichols, ABA President and CEO, to FASB Chairman Russell Golden available at `http://www.aba.com/Advocacy/LetterstoCongress/Documents/RussellGolden-FASB-011316.pdf`

set beyond that allowed under the ILM, I consider the information that managers typically use under the ILM. I construct a prediction model by removing constraints on how managers use this information and apply judgment to estimate allowances. I refer to this as the "limited information model" (LINM). I construct the limited information model in order to calibrate my estimates because one concern with the underperformance of the ILM, or with any accounting rule, is the opportunity for self-interested managers to affect its implementation. The limited information model uses the same information that is allowed in principle under the ILM, but estimates the allowance without any management judgment. If the limited information model outperforms the ILM, it would then serve as a baseline for further analysis.

It is not clear a priori how the estimates from the limited information model would perform relative to the ILM. On one hand, the ILM allowance may outperform the limited information model for at least two reasons: 1) banks possess private information, which they could use for their allowance estimation; and 2) each bank can use its own allowance model, which could improve the overall ILM estimates. On the other hand, the limited information model could outperform the ILM if the allowances from the ILM are biased due to agency problems or bounds on managers' information processing. Thus, comparing the performances of the limited information model and the ILM is ultimately an empirical question.

To preview the results from the first question (on which I elaborate below), I find the limited information model outperforms the ILM. The limited information model hence serves as a credible baseline for the analysis of my second question.

For the second research question, on the impact of expanding the information set, I build on the limited information prediction model from the first question, but extend the information set managers can use in estimating the allowance. This notion accords with the CECL, which will allow managers to use a broader set of data relative to the ILM. I refer to this model as the "full information model" (FINM). I then examine the value of incorporating broader information by comparing the performance of the expanded full information model to that of the limited information model. If there are material gains from using broader information, then I should find a significant difference between the performance of the full information model and limited information model – with the full information model having a significantly lower root mean squared error.

My empirical strategy uses quarterly data in the periods 1996 – 2012 to estimate the allowances under the limited and full information models. To answer my first research question on estimating allowances using the limited information model, I consider only historical data such as non-performing loans and loan portfolio compositions that managers use under ILM. I employ a prediction model from the machine learning literature – lasso – to predict future loan losses. The objective of lasso is to improve prediction accuracy and reduce over-fitting. It offers a feasible and objective approach to estimating allowances that could easily be implemented in practice. To predict allowances, I use a rolling-window out of sample technique, where I estimate the lasso in a particular period and then use the estimated model to predict out of sample one-quarter and two-quarters ahead. I assess if the estimated allowances from the limited information model have a lower root mean squared error than those using the reported ILM standard.

My estimation of bank loan loss allowances under the full information model hinges on my ability to proxy for the manager's information set used in the estimates. To address this challenge, I construct a dataset focused on US banks that operate in only one county. This allows me to exploit variation in county business cycles and economic conditions to construct time varying proxies for the credit risk underlying the banks' loan portfolios. My dataset includes information on current county house prices, income, and other demographic variables that I anticipate managers will use under CECL.

To empirically estimate the allowances in the full information model, I build on the limited information lasso model, but expand the information it uses to include detailed county-level economic micro data. I use the rolling-window out of sample technique for the estimates. The lasso model is particularly well suited for the full information case because it identifies relevant information that predicts loan losses and incorporates them into the allowance estimates. I compare the accuracy of the allowance estimates from the full information model to that of the allowance from the limited information model using the root mean squared error.

My findings are as follows. Regarding the first question, I find that it is indeed possible to construct a prediction model that outperforms the ILM, without expanding the information set beyond that already considered under the ILM. I find that the allowance estimated from my limited information lasso model significantly outperforms the allowance under the ILM: the limited information model accurately predicts loan losses with a lower root mean squared error (presented in top panel of Figure 1.1). The estimated allowances using the limited information model are economically meaningful and translate to an increase of 22% in allowances for the mean bank in my sample. Additionally, using the estimated allowance from the limited information model, banks would have recognized higher losses when entering the mortgage crisis beginning in 2006.

In answer to the second question, I find that expanding the information set provides no significant benefit in the model performance. The allowance estimated from the full information lasso model is directionally better than that of the limited information case, but is not substantially different (with only a slightly lower root mean squared error) from the limited information model. This finding is presented in the bottom panel of Figure 1.1. Nevertheless, the allowance estimates from the full information model do outperform the allowance under the ILM.

Finally, I provide evidence on the drivers of the performance difference between the LINM and ILM. I find that the LINM outperforms because the model assigns larger weights to key input variables relative to the ILM. In contrast, the ILM systematically under-weights the input variables. I also find that the LINM recalibrates to incorporate the underlying information. The effect of this performance difference is large for banks that operate in states with severe banking crises. In analyzing the inefficiency in the ILM and the incentives for managers to understate the losses, I find that weakly capitalized banks consistently underprovision relative to well capitalized banks.

Overall, the analysis in the paper suggests that expanding to incorporate broader information in allowance estimates is beneficial, but that these benefits primarily arise from efficient use of information, rather than by broadening the information set used. Considering that I use machine learning, what I find is that a model that uses the same information set used under ILM, but that allows no judgment to managers in estimating allowances, outperforms the ILM. These results are robust across different choice of models. When I repeat my analysis estimated using OLS, I consistently find that merely expanding the information set provides no significant benefit in

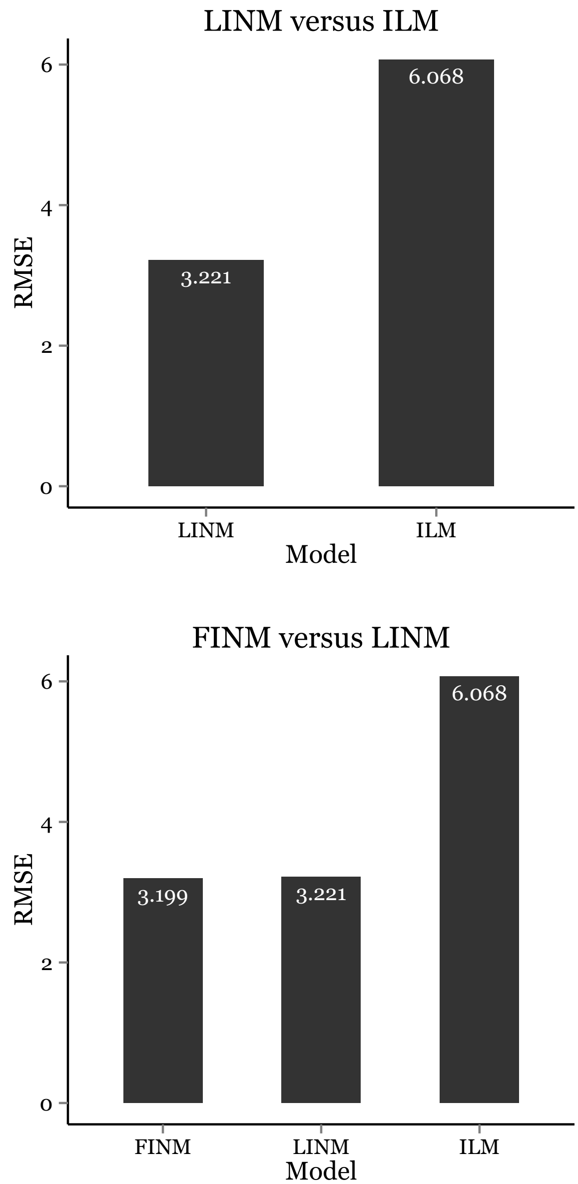


Figure 1.1: *Summary of Allowance Model Performance.*

This figure presents evidence on the performance of the model predicted allowance relative to the current GAAP's ILM in their accuracy in predicting future losses. It shows the cumulative root mean squared error for the model estimated and tested out of sample in the period 2002–2012. LINM is the root mean squared error of the predicted allowance from the lasso model that takes the limited information as input. FINM is the root mean squared error of the predicted allowance from the lasso model that takes the full information as the input. ILM is the root mean squared error of allowance under current GAAP from the financial statements. Refer Figure 1.4 for further discussion on the models.

performance accuracy. On extending the analysis to a sample of large banks operating in multiple counties, I consistently find that the limited information model continues to outperform the ILM.

The results from my paper have several important implications for bank loan loss accounting rules and their implementation. Assessing the ILM seems to suggest that even with the limited information set, using a simpler, eminently feasible, and objective approach leads to outperforming the current GAAP. One explanation for this result is the possibility that managers were reluctant to use the information to which they had access under the ILM, and that judgment was not used in an unbiased way. These findings raise the question of whether judgment and discretion enhances the quality of accounting in my setting. Finally, my paper sheds light on where the accounting rules can be improved. Despite the stress on using a broader information set in the CECL, it is not clear that information alone will improve its performance. More guidance would be helpful on how managers should implement the accounting rule.

My paper contributes to the literature in several ways. First, while a large body of accounting and banking research investigates the factors that influence bank loan loss accounting, research that examines the role of the accounting standards is scarce (e.g., Beck and Narayanamoorthy, 2013). A unique feature of my paper is that it builds on the institutional details around bank loan losses, and focuses on the implementation aspects of the accounting rules. Second, I identify a setting in the CECL rule to visit the longstanding question over discretion versus rules in financial reporting. I provide evidence building from the theoretical notion of loan loss allowances, and develop a prediction model. I consider an objective approach from machine learning to estimate the allowances. This aids in better understanding the implications of bank managers exercising more discretion in applying accounting rules. Third, I propose a way to implement the CECL rule by using objective and verifiable information, thereby examining aspects of the model before it goes into effect in December 2019. Finally, the machine learning techniques used in the paper open up opportunities for further research in accounting, particularly in scenarios where the objective is to build a prediction model and where causal inference is not the primary aim.

The rest of the study proceeds as follows. Section 1.2 discusses the motivation and background of the paper, while Section 1.3 lays out the empirical framework. Section 1.4 discusses details of the data. Section 1.5 provides the empirical results of my analysis and Section 1.6 discusses possible interpretations of the results. Section 1.7 concludes.

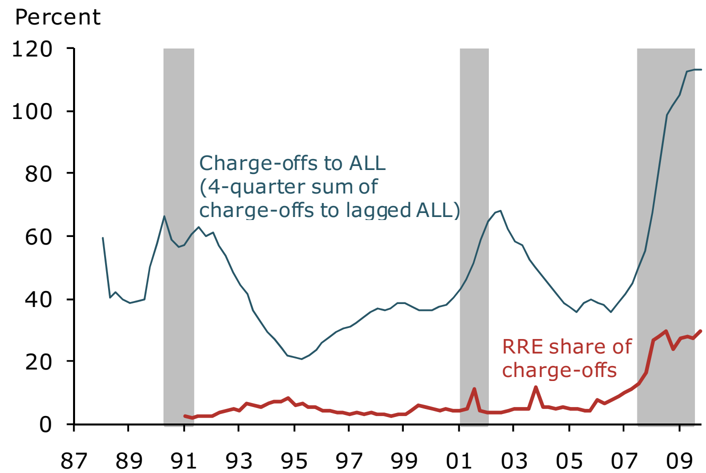
1.2 Motivation and Background

Loans held to maturity on banks' balance sheets are a gross asset that reports the remaining contractual principal on loans in the portfolio. The allowance for loan losses, a contra-asset account, has several components, but the largest consists of allowances for loans collectively evaluated for impairment that fall under FAS 5.[3] The current accounting rules from the FASB require banks to assess whether there is any objective evidence of a "loss event" that indicates that a loan or group of loans is impaired. If there is objective evidence of an impairment (probable condition), the amount of the loss incurred needs to be estimated (measurable condition). This methodology is referred to as the incurred loss model (ILM).

The notion of *incurred* is an application of the general requirement in GAAP that an accounting event (like loan delinquency) must occur for accounting recognition to occur. Therefore, under ILM, banks can recognize losses for an impaired asset only when a loss is probable based on past events. The rationale for these requirements is that it limits managers' ability to manipulate financial statements. Thus, FAS 5 prohibits firms from accruing for losses that are not currently incurred, even if those losses are expected or from possible future events (see Ryan, 2012).[4] Historically, the issue whether allowances should be provided for potential or possible future losses, has received the most attention from the SEC and bank regulators having issued interpretive guidance and various publications related to the application of the ILM (Baskin, 1992; SEC, 1997b). The application hinges on two key guidance releases, FRR28 and SAB102, both of which are

[3]FAS 5 refers to the original FASB pronouncement FAS 5, Accounting for Contingencies, which is included in the FASB Accounting Standards Codification (ASC) subtopic 450–20, Contingencies: Loss Contingencies. The other two components are: FAS 114 (Also known as ASC 310–10–35) for loans individually determined to be impaired and SOP 03–3 (ASC 310–30) for loans acquired with deteriorated credit.

[4]Also, FAS 5 Paragraph 30 states that "... exposure to risks does not mean that an asset has been impaired or a liability has been incurred. The condition for accrual ... is not met with respect to loss ... that may occur after the date of an enterprise's financial statements. Losses of those types do not relate to the current or a prior period but rather to the future period in which they occur."



Source: Haver Analytics.
Note: Quarterly data. ALL: allowances for loan losses.

Figure 1.2: *Ratio of Chargeoff to Allowance*

This figure presents ratio of charge-offs to loan loss allowances for all loans, and the share of residential real estate charge-offs at commercial banks as adapted from Furlong and Knight (2010).

faithful to the FAS 5 loss condition.[5]

During and after the financial crisis, the concerns resurfaced that this aspect of GAAP did not allow banks to accrue adequate reserves, which delayed the reporting of losses and hence exacerbated the severity of economic downturns (Laeven and Majnoni, 2003; Dugan, 2009; Beatty and Liao, 2011). The criticism of the model was in limiting banks from recognizing loan losses that were expected, but did not yet meet the probable threshold. Indeed, in past downturns, realized loan losses were not adequately reflected in banks' provisions. And using historical loss rates for incurred losses, which were low in pre-crisis years, underestimated the provisions. This issue is illustrated in Figure 1.2, which shows that realized loan losses rose relative to lagged banks' provisions consistently at the beginning of any given period, and skyrocketed during the 2008 – 2009 mortgage crisis.

The ILM with its focus on using historical information for loan loss recognition is inherently backward-looking. Thus, banks are allowed to use only a subset of the available information to predict loan losses for the purpose of estimating allowances. To illustrate, consider a bank with a

[5]See further discussions in Ryan and Keeley (2013).

portfolio of mortgage loans that observes home price declines. Under FAS 5, the bank recognizes losses based only on loans that are currently delinquent when estimated collectively. But one can anticipate larger losses if the prediction is further conditioned on current housing market shocks that will yield future defaults, and if all available information are used in the estimation. This constraint potentially impedes banks from setting adequate levels of reserves, misstating the value of the assets, and hindering transparent reporting.

In discussing the severity of the issue in one of his addresses, Ben Bernanke, then Chairman of the Board of Governors of the Federal Reserve, stressed that "there is considerable uncertainty regarding the appropriate levels of loan loss reserves over the cycle. As a result, further review of accounting standards governing ... loan loss provisioning would be useful, and might result in modifications to the accounting rules that reduce their pro-cyclical effects without compromising the goals of disclosure and transparency" (Bernanke, 2009).

The procyclicality referred to by Bernanke is the consequence of the ILM preventing banks from recording adequate provisions during favorable periods in the economic cycle. As a result, incurred losses are lower than the long run average i.e., across the cycle. When the credit quality of bank loans portfolios deteriorates during downturns, banks must accrue excessive provisions, potentially magnifying the impact of the economic cycle on the loan loss allowance by reducing the bank's regulatory capital at times when it is expensive to raise capital (FSB, 2009).[6]

Bernanke's sentiment was echoed by the US GAO, who in their report and testimony to Congress examining bank failures (see GAO-US, 2013a,b) stated that the "Federal banking regulators have also noted that requiring management at failed banks to recognize loan losses earlier could have helped stem losses" and "potentially lessened the impact of the crisis, when banks had to recognize the losses through a sudden series of provisions to the loan loss allowance, thus reducing earnings and regulatory capital." The GAO also expressed these concerns about community banks, which are concentrated in their lending, and are hence more exposed to local shocks.

[6]The world leaders during the G20 Washington Summit in their response to the crisis called for the development of "The IMF, expanded FSF, and other regulators and bodies should develop recommendations to mitigate procyclicality, including the review of how valuation and leverage, bank capital, executive compensation, and provisioning practices may exacerbate cyclical trends." (G20 Summit, 2008)

These events together have made FASB and IASB reconsider the loan loss provision accounting rules to address the criticism. The FASB in its response issued a proposal for public comment for an alternative loan loss provisioning rule that would accommodate the use of more forward-looking information and stress on expected losses. After several years of deliberation, FASB issued a final standard as of June 2016 – the CECL. The IASB separately replaced IAS 39 with the guidance IFRS 9 for Financial Instruments, which is also based on an expected credit loss model, issuing a final standard in July 2014.[7]

1.2.1 Current GAAP Rules – Implemention

The incurred loss model focuses on banks' discretion in identifying probable and quantifying estimable losses. FAS 5 defines "probable" as "the event or events are likely to occur" (paragraph 3). Thus, the model responds to an event that triggers the recognition of loan losses. This aspect of the rule was further stressed by Thomas Curry, the head of the OCC while addressing the AICPA conference in 2013 (see OCC, 2013): "As you know, current accounting rules prevent banks from provisioning for an impaired asset until a 'triggering event' occurs. In other words, banks must wait until the event has already occurred to recognize the loss."

In practice, the threshold for probable is defined as 70% or more likely (see Ryan, 2012). When loans are underwritten, banks classify them into pools of loans with similar risk categories.[8] Once the loans are classified, the allowances are estimated by identifying trigger events for losses based on the severity of delinquency and payment status. The banks primarily use information on past due loans such as 30 days past due, 90 days past due, and non-accrual along with historic charge-off rates for the estimation (Balla and McKenna, 2009).[9] Banks also consider the composition of loans in their portfolio in determining allowances. Thus, historical information is used as basis for

[7]Although the impairment project began as a join effort of the FASB and IASB, the constituent feedback on the boards dual measurement approach led the FASB to develop its own impairment model. The IASB continued in developing the dual measurement approach and issued the guidance that was amended to IFRS 9 (see Deloitte, 2015).

[8]These risk categories vary by banks. They fall broadly under pass, special mention, substandard, doubtful, or loss. The loans can move between loan categories during its term in the portfolio.

[9]See (Walter, 1991) for further explanation on how banks categorize defaults and estimate allowance.

estimating the incurred losses.

To summarize, with the stress on the probable threshold, and the focus on using historical information to predict future losses, the ILM allows banks to use only a limited set of information that are available to them. The rule limits discretion, by prohibiting banks from accruing for loans that are not currently incurred, even if they are a result of predictable or possible (less likely than not) future events.[10,11]

1.2.2 CECL Rule and the Use of Broader Information

The main objective of the CECL is to enhance transparency in reporting expected future loan losses, and therefore capture current risks in the banks' portfolio. It aims to achieve this objective by eliminating the "probable" threshold that exists under the ILM. Therefore, a trigger event would no longer be required to estimate the loan losses. It would also weaken the incurred condition that exists under the ILM (FASB, 2016; Acharya and Ryan, 2016).

Furthermore, the CECL rule alters the information set banks can use in estimating future expected losses, compared to the ILM. The rule broadens the information considered when measuring credit losses to include forward-looking information. Therefore, the CECL focuses on allowing managers to use a broader set of information about past events, current conditions, and "reasonable and supportable forecasts" relevant to assessing future credit losses.

The CECL model reflects a fundamental shift in estimating allowance compared to the incurred loss model. In contrast to the ILM, which requires recognizing losses only on delinquent loans, the CECL requires recognizing losses on all loans (including the ones that are in current standing). The CECL, in principle, should provide early warning signs of deterioration in the economic conditions relative to the ILM. Summarizing, the CECL allows managers to incorporate the full set of information that are available to them in estimating the allowances.

[10] Paragraph 59 of FAS 5 of states that that loan-loss provisions should reflect events occurring within the reporting period and not anticipate future events: "Further, even losses that are reasonably estimable should not be accrued if it is not probable that an asset has been impaired or a liability has been incurred at the date of an enterprise's financial statements because those losses relate to a future period rather than the current or a prior period. Attribution of a loss to events of the current or prior periods is an element of asset impairment or liability incurrence."

[11] See also Footnote 4.

1.2.3 Evaluating the ILM and the CECL

My empirical focus is on examining the ILM and the CECL rules by developing prediction models for future loan losses to estimate allowances. I exploit the differences in the information sets allowed under these rules for the estimation. The current ILM rule not only limits the information managers can use, but also the discretion in constraining *how* this information is used. Therefore, my investigation of the models involves considering (i) an objective approach, which would remove constraints on how the information is used, thus removing the discretion and (ii) the information set the model has access to.

This discussion motivates the principal empirical tests based on two models that I estimate. First, I ask whether it is possible to construct a predictor of future losses that would outperform the ILM approach, without having to expand the information set (limited information model). Second, I assess the impact of expanding the information set to the limited information model by including the broader set of data as under the CECL rule (full information model).

Assessing both the limited information model and full information model are necessary to answer my research question and to understand the implications of the ILM and CECL models. Documenting evidence of any performance difference between the full information model (i.e., my implementation of the CECL) and the ILM does not address whether the difference is from the better use of information, as captured by the model, or from the broader set of information used. On the other hand, the limited information model provides a set of baseline results that can be used to gauge its performance relative to both the ILM and the full information model.

There is, however, a key limitation in evaluating the limited information and full information models. The allowances under CECL are not observed in the bank financial statements since the rule change goes into effect in the fiscal year beginning December 15, 2019. Consequently, I offer a new approach to estimating counterfactual allowances under the two models, which is based on mimicking the managers' decision-making and performing out of sample predictions. The innovation lies in treating this as a prediction problem, rather than a problem of causal inference. Machine learning methods are better suited than traditional regression approaches for this purpose. I use an objective approach, lasso, which is widely used in the statistical and machine learning literature.[12] I discuss the lasso methodology in Section 1.3.4, and its implementation relevant to this paper in Section 1.5.1.

1.2.4 Research on Loan Loss Accounting

This study is related to several lines of research in accounting and banking. First, it contributes to the large body of research on the factors that influence bank loan loss accounting. Prior research documents the incentives of bank managers to smooth earnings, circumvent capital requirements, and reduce taxes (Ahmed *et al.*, 1999; Wall and Koch, 2000; Beatty *et al.*, 2002; Collins *et al.*, 1995). But as argued in Ryan (2012), the results from this literature have been overall inconsistent. In my paper, I offer another view to the literature on the role of accounting standards, and also emphasize that banks' loan loss estimates also depend on FAS 5 ILM. Indeed, as Dechow *et al.* (2010) state, a general problem accounting researchers encounter in identifying income smoothing is the difficulty in disentangling (i) the fundamental earnings process (ii) the accounting rules and (iii) intentional earnings manipulation. A key aspect of my analysis is that I do not explicitly model agency problems, and the performance difference between my models and the ILM arises primarily due to the accounting rules.[13]

Second, my work relates to the debate on the relationship between transparency and stability. This literature focuses on whether increasing transparency enhances or impairs stability. The arguments for the positive effects of transparency are that it enhances market discipline, increases timely regulatory intervention, and mitigates panics (Rochet, 1992; Ratnovski, 2013; Granja, 2013).[14] The opposing view argues that transparency can lead to bank runs that are driven by coordination failures, and can also lead to inefficient investment decisions (Morris and Shin, 2002; Goldstein and Sapra, 2014), while Dang *et al.* (2014) argue that bank opacity avoids adverse selection in the market, allowing the deposit mechanism to work better. There is a consensus, however, that one of the primary channels by which bank financial reports can affect stability is by

[12] Note that the name "lasso" is an acronym for Least Absolute Selection and Shrinkage Operator.

[13] This modeling assumption that there are no agency problems is primarily for convenience and tractability.

[14] See also Bushman (2015) for a discussion of the topic.

increasing public information about the risk and economic conditions to which banks are exposed (Acharya and Ryan, 2016). Loan loss allowance estimates reflect a fundamental aspect of the risk attributes of a loan portfolio. My study considers a specific setting to discuss how the new rule, whose objective is to enhance transparency, is likely to work in practice.

Third, this paper contributes to the policy discussion around accounting standard setting for loan loss provisions.[15] These discussions have been motivated by the need to use more forward-looking measures to mitigate the limitations of the ILM. Studies on forward-looking measures include Bushman and Williams (2012) that use a cross-country sample to argue that forward-looking provisions, if designed to smooth earnings increase risk taking in banks.[16] In modeling expected losses, Harris *et al.* (2017) develop a measure to capture one-year forward expected losses. They used cross-sectional regression analysis and find the measure performs well in predicting next year losses and contains incremental information relative to fair value disclosures in explaining future losses. The case of Spain provides useful data for studying the issue, as Spanish banks have employed an alternative approach in dynamic provisioning for loan losses since 2000 to reduce pro-cyclicality (see Fernández *et al.*, 2000; Saurina, 2009a,b). Using this setting, Jiménez *et al.* (2012) examine the changes in credit availability after dynamic provisioning went into effect. The authors find that the growth in credit during upturns and credit supply contraction during downturns are significantly reduced after the implementation of the dynamic provisioning approach. Similarly, Pérez *et al.* (2011) study income-smoothing in Spanish banks around the implementation of dynamic provisioning, and find the effect only from 1988 to 1999. They find that between 2000 to 2004 when banks used dynamic provisions, transparency increased, but there was no evidence of income-smoothing. I add to this discussion by performing an empirical study to understand how forward-looking measures would have performed in a US setting.

Finally, my paper designs and implements a new approach to estimate allowances that incorporates more current information, as would be observed under CECL. To construct the information

[15] See further discussion by Barth (2007) on the case for researchers to develop studies relevant to issues faced by the standard setters.

[16] Consistent with these results, Bushman and Williams (2015) find that U.S. banks that tend to delay loss provisioning are associated with contractions in their balance sheet, and contribute to systemic risk during periods of downturns.

set for my analysis, I use detailed micro data to empirically build proxies that capture the credit risk in loan portfolios. To identify the credit risk, I build on seminal work by Mian and Sufi (2009, 2011) who study the causes and consequences of the credit crisis by examining the role of household debt and defaults in the period. The implications to bank accounting arise from the fact that household debts are typically assets on banks' balance sheets. Thus, it is reasonable to posit that the factors that drive household debt defaults would be the information that managers would consider to predict future losses.

In implementing my estimation, I build on emerging research by economists (Mullainathan, 2014; Kleinberg *et al.*, 2015) who argue for the value of using different empirical tools from ML that are designed primarily for prediction. The authors give examples of applying the tools to policy questions, which have prediction as an objective of interest. In estimating the allowances using broader information, I approach my question as a case of bank managerial decision making and model it using machine learning (ML) techniques to predict future losses. This idea is a natural extension of prior accounting studies that have focused on the topic of using ML (referred to as an expert system or artificial intelligence) for human information processing and decision making.[17]

1.3 Empirical Framework

This section develops the statistical framework for predicting allowances in the limited information and full information model. I use this framework to motivate the empirical analysis, and gradually discuss the various ingredients of the prediction technique. My framework builds on modeling the bank manager's decision by considering 1) the information she has access to and 2) the information that she is allowed to use given the rule. The assumption is that a reasonable manager uses the information set and the realized losses in prior periods to build a prediction model. The model would then be used to predict losses out of sample. If the bank maintains the allowance at this

[17]Some of these studies have used the lens model analysis (Zimmer, 1980; Kim and McLeod Jr, 1999). For other examples, see work by Libby (1976), Casey (1983) and Wright and Willingham (1997). In other works, Mengle (1990) discusses about the efforts from the accounting industry developing "expert systems" to capture the estimation of the quality of the loans and ascertaining the market values of loan loss reserves.

estimate, then the total loans, less the allowance, would be the best estimate of the collectible value of the loan at the evaluation date.

My empirical strategy in the paper follows a two-step procedure to estimate the allowance under the limited and full information models. I first use the information set allowed under the relevant rule, and build a prediction model of loan losses in an estimation period. The model is estimated every quarter, where I follow a rolling-window technique with the estimation period expanded every quarter to use information only up to that quarter. The rolling-window technique helps to avoid any hindsight bias. In the second step, I use the model developed in the estimation period to make predictions in the subsequent time period, the out of sample test period. The predictions in the test periods are then used to evaluate the model's performance. This algorithm is further discussed in Section 1.5.1.

Let the information set a manager has access to be denoted by $\mathscr{I}_t$, which could be used at time t to predict losses in estimating the bank's loan loss allowance LLA_t. The $\mathscr{I}_t$ refers to the vector

$$\mathscr{I}_t := \left[\{X_0, X_1, \ldots, X_t\}, \{M_0, M_1, \ldots, M_t\}\right],$$

where X captures bank-level variables, such as the portfolio loan balance, information on the delinquency status of the loans, net-chargeoffs. The set M contains macroeconomic indicators that affect the loan performance of the bank such as home prices, unemployment rates, and wages. Thus, a subset of $\mathscr{I}_t$ is the information banks can use under ILM, denoted by $\mathscr{I}^{ilm}$. The set $\mathscr{I}_t$ also contains the broader information set allowed under CECL, which is denoted by $\mathscr{I}^{cecl}$. The $\mathscr{I}^{ilm}$ is hence the limited information set containing variables in X, while $\mathscr{I}^{cecl}$ is the full information set containing X and M.

I first estimate allowances using the limited information model. I then expand the information the limited information model uses for the full information model. The accuracy of each model is assessed by how well the estimated allowance captures actual net-chargeoffs. This idea is consistent with the accounting identity:

$$\begin{aligned} \mathrm{LLA}_t &= \mathrm{LLA}_{t-1} + \mathrm{RECOV}_t - \mathrm{CO}_t + \mathrm{LLP}_t \\ &= \mathrm{LLA}_{t-1} - \mathrm{NCO}_t + \mathrm{LLP}_t \end{aligned} \tag{1.1}$$

where LLA_t is allowance for loan losses at end of t, RECOV_t the amounts that have previously been charged off but are recovered during this time period, CO_t the gross amount of all loans charged off against the LLA losses. The income statement effect is captured by LLP, the loan loss provisions. Thus, the precision of the LLA_{t-1} is assessed by how well it predicts future net-chargeoffs (NCO) at t.

1.3.1 Current US GAAP

I begin by studying the banks' loan loss estimation strategy using the ILM, which forms the motivation for my limited information model predictions. For the discussions around the relevant statements of the rule, refer to Section A.1. At time t, the LLA_t is estimated by predicting CO_{t+1}, using information from t, based on the subset of data that could be used i.e., $\mathscr{I}_t^{ilm}$. This timing notation captures the aspect under the ILM, even though the manager has access to larger information from t, only is allowed to use a subset in $\mathscr{I}_t^{ilm}$.In the discussions that follow, I use charge-offs to motivate the prediction models but my actual empirical estimates uses the net-chargeoffs. I initially assume that the manager predicts CO one period ahead, such that estimates LLA_t to cover CO_{t+1}. This assumption is relaxed to consider multi-period losses later in this section. Formally,

$$\mathrm{LLA}_t = \mathbb{E}[\mathrm{CO}_{t+1} | \mathscr{I}_t^{ilm}]. \tag{1.2}$$

At the evaluation date t, the manager uses historical data in the periods $\tau = 0, \ldots, t$ and uses information on delinquencies and nonperforming loans and their corresponding charge-offs to build a prediction model of CO. She then uses the model estimates to make out of sample predictions, with a new realization of information at t that she can use. If the bank manager uses an OLS model, then

$$\mathbb{E}[\mathrm{CO}_{t+1}|\mathscr{I}_t^{ilm}] = \mathbb{E}[\alpha + \beta * X_t + \epsilon|\mathscr{I}_t^{ilm}], \tag{1.3}$$
$$\mathrm{CO}_{t+1} = \alpha + \beta * X_t + \epsilon.$$

Therefore, it follows that at time t

$$CO_t = \alpha + \beta * X_{t-1} + \epsilon, \tag{1.4}$$

where CO_t is the realized credit losses of the bank at t.

Empirically, a researcher regresses CO_t on X, where X is bank-level covariates defined earlier, and recover the estimates $\hat{\alpha}$ and $\hat{\beta}$. Substituting the parameter estimates in 1.2 and making predictions with the new realization X_t,

$$\begin{aligned}\hat{\alpha} + \hat{\beta} * X_t &= \widehat{CO_{t+1}} \\ &= \widehat{LLA}_{t(ilm)}.\end{aligned}$$

This forms the basis for estimating allowances under the limited information model. I use the lasso approach for the predicting the allowances, which serves as my baseline to compare it performance to the allowance from the ILM that are reported in the bank financial statements. I attribute the difference in performance to how the limited information model uses the information set while estimating the allowance, and attribute the effect to discretion in the ILM.

1.3.2 CECL Model

This section forms the motivation for my estimation in the full information model. The key feature of the expected loss model is that it alters the information set that banks can use while estimating their LLA, without having to wait for a triggering event.

Under the CECL, bank managers predict CO using a broader set of information that reflects current condition, and thus any deterioration of credit. Formally:

$$\mathrm{LLA}_{t(cecl)} = \mathbb{E}[\mathrm{CO}_{t+1}|\mathscr{I}_t^{cecl}] \tag{1.5}$$

over $\tau = 0, \ldots, t$. The manager as of evaluation date t incorporates all information up to that point in time. Similar to the derivation in Equation 1.3,

$$CO_t = \alpha_{cecl} + \beta_{cecl} * X_{t-1} + \gamma_{cecl} * M_{t-1} + \epsilon, \tag{1.6}$$

i.e., the bank predicts CO based on its information at t. This idea is empirically equivalent to regressing CO_t on X_{t-1}, but also includes county level economic information M. I then use the parameters from the predicted model, $\hat{\alpha}_{cecl}$ and $\hat{\beta}_{cecl}$, to estimate at period t

$$\begin{aligned} \hat{\alpha}_{cecl} + \hat{\beta}_{cecl} * X_t + \hat{\gamma}_{cecl} * M_t &= LLA_{t(cecl)} \\ &= \widehat{LLA}_{t(cecl)}. \end{aligned} \tag{1.7}$$

The $\widehat{LLA}_{t(cecl)}$ is my estimated allowance under full information model. The prediction model I use is the lasso. I interpret the analysis of the limited information model and the full information model under the maintained hypothesis that these estimated measures reflect an unbiased and knowledgeable party's estimate of future losses using the information. I attribute the difference in the performance of the limited information and full information model to the broader information incorporated in the full information model.

1.3.3 Measuring Loan Losses

In the above model, I assume the manager's estimate of LLA at time t absorbs loan losses CO in future periods $t+1$. In reality, the bank estimates LLA in 1.2 and 1.5 by developing a model that predicts credit losses in multiple subsequent periods towards the maturity of the loans.

To expand the model's loss horizon, let the adequate level of LLA the bank estimates at the end of a given quarter be the amount needed to cover the projected loan losses over the next four quarters. At time t,

$$g = \sum_{\tau=t+1}^{t+4} CO_\tau + NPL_{t+4}.$$

This assumption of adequate allowance is similar to the requirement in OCC (1998).[18] To

[18]OCC (1998) states that "Many banks consider coverage of one year's losses an appropriate benchmark of an

capture a longer horizon, following Fillat and Montoriol-Garriga (2010) I include non-performing loans at the end of the four quarters. This broader measure of loan losses, the dependent variable in the lasso analysis, is the sum of rolling four quarter net chargeoffs, and loans ninety-days past due and non-accrual at the end of the rolling-window's fourth quarter. This measure avoids any double-counting when non-performing loans become charged-offs in the future periods, as the ninety-days past due and the non-accrual are measured at the end of the fourth-quarter in the rolling window. This approach provides a more conservative estimate than what CECL will require.

1.3.4 Lasso Approach

The above discussion highlights the importance of prediction in modeling the manager's estimates of the loan loss allowance. I thus build on tools from the machine learning literature for estimating predictions in my study. The reasons for this approach are as follows: (i) The tools, including OLS, designed for causal inference are meant to generate an unbiased estimate of $\hat{\beta}$ when used in sample.[19] They do not usually yield the most accurate prediction $\widehat{Y}$. (ii) As a researcher, the manager's mapping of the information set in predicting losses to estimate the allowance is unobservable. Managers would implicitly develop a model using the data they have and make predictions out of sample. Machine learning techniques are designed primarily for just this purpose: that of improving accuracy while making out of sample predictions (For further discussion of using ML methods in empirical social science research, see Mullainathan (2014); Einav and Levin (2014); Athey and Imbens (2015)).

I build on the framework provided in the recent economics literature by Kleinberg *et al.* (2015) for using tools from ML for prediction problems, particularly in empirical policy research. The objective of the ML tools is to reduce prediction errors that are a combination of bias and variance.

---

adequate reserve for most pools of loans. ...A one-year coverage period is generally considered appropriate because the probable loss on any given loan in a pool should ordinarily become apparent in that time frame." Further, the Federal Reserve as part of the stress test argues that "The appropriate level of ALLL at the end of a given quarter is generally assumed to be the amount needed to cover projected loan losses over the next four quarters." http://www.federalreserve.gov/bankinforeg/stress-tests/2015-Appendix-B.htm

[19]A key difference is that models that are used for prediction purposes are not intended for hypothesis testing in the traditional sense.

In contrast, the focus of OLS is to get unbiased estimates by minimizing the squared error, i.e.,

$$\min_{\beta \in \mathbb{R}^p} \left\{ \sum_{i=1}^{N} (Y_i - x_i^\top \beta)^2 \right\}.$$

By ensuring zero bias, OLS allows no trade-off with variance and so minimizes in-sample error. But for prediction, the primary objective, as opposed to fit in sample, is the accuracy out-of-sample. Ensuring zero bias in sample constrains the performance of the prediction model out of sample. The ML techniques (see Hastie *et al.*, 2009) were developed mainly to empirically handle this bias-variance trade-off and maximize prediction performance. I thus use the lasso predictor (Tibshirani, 1996), which is among the most widely used models under the class of regularization methods.

The lasso estimator minimizes:

$$\min_{\beta \in \mathbb{R}^p} \left\{ \sum_{i=1}^{N} (Y_i - x_i^\top \beta)^2 + \lambda \sum_{j=1}^{p} |\beta_j| \right\}.$$

This estimator is an extension to OLS. However, the second term, which is a function of λ and β called a shrinkage penalty, is small when some of the β coefficients are close to zero. The $\sum_{j=1}^{p} |\beta_j|$ term has the effect of forcing some of the estimated β coefficients to be exactly equal to zero when the λ is sufficiently large (see further discussion in Tibshirani (1996)).[20]

The key idea is that the parameter λ can be chosen using the data. For this choice, I employ a commonly used cross-validation technique. The algorithm performs the bias-variance trade-off by systematically picking a λ that performs well out of sample, reducing the problem of overfitting. The key is that using lasso allows me to fit a model containing a large number of predictors with some of the coefficients that are regularized (shrunk) towards zero.

As discussed in Kleinberg *et al.* (2015), ML techniques are an extension of non-parametric statistics. Specific to my research question, they provide a disciplined way to pick a model that would approximate the manager's function in mapping the information set to predict losses CO_t, while estimating allowances. The lasso offers an eminently feasible, simple, objective approach to implement the accounting rules. The lasso performs variable selection to include predictors

[20]Note that OLS is a special case where $\lambda = 0$, that is there is an infinite price on bias relative to variance.

that are a subset of the variables. They also allow for developing a model searching over a set of variables and functional forms.

1.3.5 Predicting Loan Losses

Here I outline the functional form for predicting loan losses to model charge-offs CO in equations 1.2 and 1.5. Consider the future expected chargeoffs,

$$\mathbb{E}(\text{CO}) = \text{loan balance * pd * lgd},$$

where pd is the probability of default, and lgd is the percent-loss given default. The term pd * lgd can be collapsed into historical charge-off rate (COrate). Estimating the expected loss is equivalent to predicting the charge-off rate, which is equal to:

$$= \text{loan balance}_{t-1} * \widehat{COrate}_t \tag{1.8}$$

The $\widehat{\text{COrate}}_t$ can be predicted from historical COrate_{t-1}, adjusting to incorporate new information for subsequent defaults. To illustrate, consider the unemployment rate at the county level, and then rewrite the estimate in a regression framework for a bank i that operates in county j in time t. In the actual empirical model used in the paper, I use a comprehensive set of macroeconomic indicators.

$$\begin{aligned} \text{COrate}_{ijt} &= \beta_1 * \text{COrate}_{it-1} + \beta_2 * \text{NPL}_{it-1} + \beta_3 * \Delta\text{unemprate}_{jt-1} + \\ &\quad \beta_3 * \mathbf{X}_{it-1} + \beta_4 * \mathbf{M}_{jt-1} + \text{fixed effects} + u_{ijt}. \end{aligned} \tag{1.9}$$

Refer to Appendix Section A.3 for further discussion of the empirical model. I use the specification from 1.9 to predict credit losses for my analysis with the time period t in quarters. This forms the basis for the variables used in my lasso model. The dependent variable I use is the sum of four quarter net charge offs and nonperforming loans. I use the predicted losses from the estimation as my $\widehat{LLA}_{t(cecl)}$. The data I use for the estimation comes from a cross sectional panel of banks with quarterly loan loss data between 2002Q1 and 2012Q4.

1.4 Data

To estimate allowances in the full information model hinges on my ability to observe the information set of the manager that predicts the credit risk in the banks' portfolio. As a result, I restrict the sample to banks that primarily do their lending in one county in the US for my analysis. I then obtain county data, such as housing prices, unemployment rates. The advantage of this research design is that it not only allows me to precisely identify the economic signals to predict future losses, but also to use a consistent sample for my limited information and full information models. Table A.2 in the appendix provides definitions of all variables used in the paper.

1.4.1 Sample selection

I use the Summary of Deposit (SOD) statement from the FDIC to identify banks that operate in only one county. The SOD is the annual survey of branch office deposits for all FDIC-insured institutions. All institutions with branch offices are required to submit the survey, but institutions with only a main office are exempt. The data includes exact address of the branch offices, including the county. I aggregate the underlying branch data to the bank-county level and identify banks that have all their operations in the same county. The implicit assumption is that these banks do lending in the same county.[21]

To avoid misclassification, I merge the bank-county data with the list of banks the FDIC has designated as community banks for research purposes (FDIC, 2012). The FDIC's research definition of a community bank fits my analysis well for two reasons. First, these banks have lending and deposit taking as their main business focus, and second, their business is fairly circumscribed in a geographic area. These aspects tie both bank characteristics, as well as the demographics of

[21] My conversations with various bank regulators strengthen the reasonableness of this assumption. The SOD gives only deposit data at the branch level. I use deposits, aggregating them at the county level, to capture the geographic concentration of bank lending and to exploit the variation across the bank's loan portfolio. This assumption is reasonable as under section 109, banks are prohibited from opening, or acquiring branches outside their home state primarily for deposits. This limitation is under the Community Reinvestment Act (CRA) that was enacted by the Congress to ensure that the bank branches do not take deposits from a community without lending to the community (FDIC, 2014).

the borrowers from these banks. This identification hinges on the assumption that both the bank (loans) and the borrowers are exposed to the same economic shocks at the county level. This level of abstraction allows me to model information that is not available in the call reports of the banks and to infer the quality of loans in the bank's portfolios.

Using the above criteria generates a sample of $306,666$ bank-county-quarter observations, for $7,086$ banks that operated in 1996 to $3,170$ banks that operated in 2012 as shown in Table 1.1. Given the stringent restriction on one-county lending, this represents a significant number of banks for analysis and reflects a sizable economic activity. Together these banks represent 59% of all commercial banks in 1996 in number, decreasing to 43% of all banks in 2012. They operated in $2,337$ $(1,519)$ counties in 1996 (2012). The US map in Figure 1.3 shows the distribution of the counties for 2002Q1 with banks that operate only in that county. Sample banks have a presence across the US, and are particularly concentrated in the Midwest, Northeast, mid-Atlantic, California, eastern Texas, and Washington. In 2002, the mean number of concentrated banks in the counties in which they operated was 2.56, while this number was 2.08 in 2012. Together, this information shows the diversity of the commercial banking in the US that allows me to run this quasi-experiment in exploiting economic cycles across counties.

1.4.2 County Data

The main sample I construct is from 1996 – 2012 at the bank–county–quarter level. For these counties, I follow work by Mian and Sufi (2009, 2011) on building micro data on local economic shocks.[22] I obtain monthly housing prices at the county level from a proprietary data set from Core Logic and supplement it with data from the Federal House Financing Agency (FHFA).[23] I use these data together to construct quarterly shocks to housing market performance in these counties. The quarterly county-level unemployment rates are from the Bureau of Labor Statistics.

I also collect average adjusted gross income at the county level from the IRS. This data gives

[22] The authors in their work argue the importance of using county variation and encourage using micro data for analysis.

[23] Both these sources use the same house repeat sales data to construct the home price index and are among the some of the best data sources on home prices (see Mian *et al.*, 2013).

Table 1.1: *Selection of Sample Banks*

Year	All US Banks	Number of Single-County Banks	Number of Counties
1996	12,047	7,086	2,337
1997	11,600	6,597	2,270
1998	11,079	6,081	2,199
1999	10,699	5,727	2,140
2000	10,430	5,500	2,089
2001	10,031	5,255	2,021
2002	9,690	5,031	1,966
2003	9,505	4,763	1,911
2004	9,313	4,599	1,874
2005	9,174	4,347	1,806
2006	9,033	4,148	1,743
2007	8,877	3,989	1,706
2008	8,635	3,828	1,658
2009	8,342	3,678	1,608
2010	8,018	3,518	1,577
2011	7,650	3,342	1,546
2012	7,377	3,170	1,519

This table shows yearly distribution of number of sample banks used in this study. The second column shows the number of all commercial banks regulated under FDIC. The third column shows the number of banks having hundred percent of their deposits in one county. The third column shows the number of different counties that these banks operate in. The sample is formed by merging the FDIC summary of deposits data along with the list of banks that are designated as community banks by the FDIC in the period.

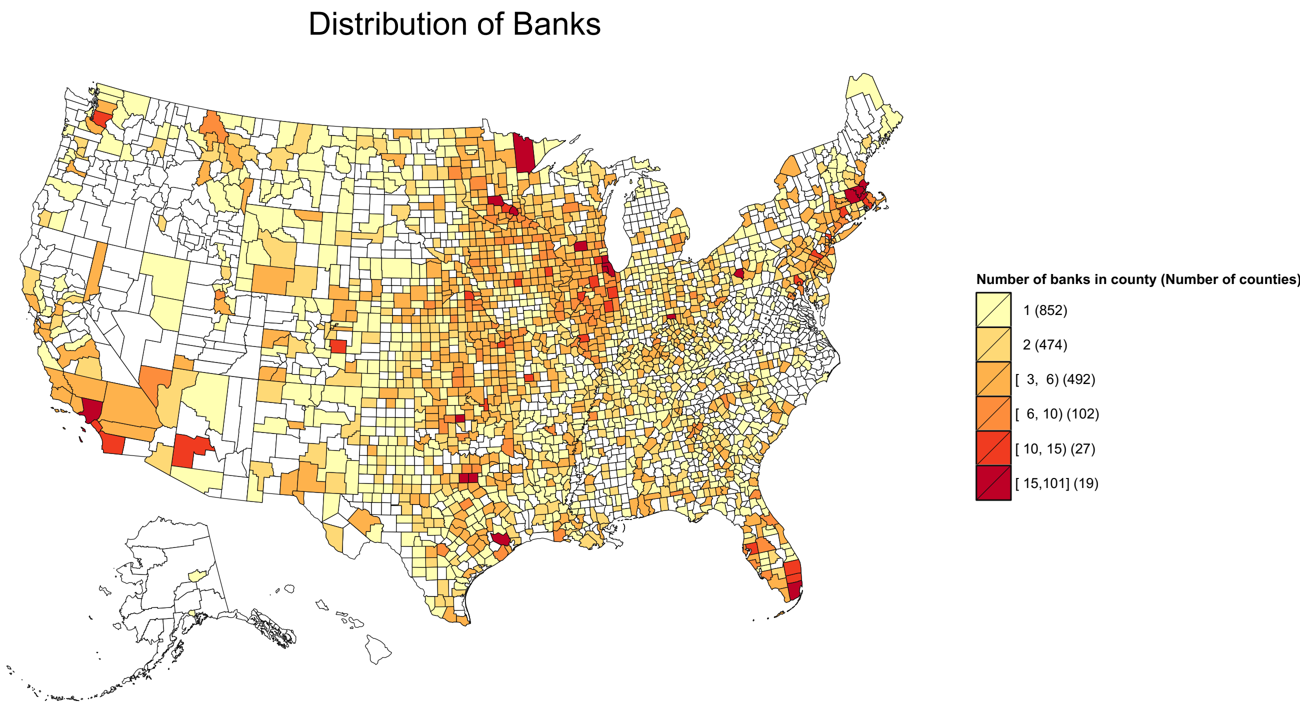


Figure 1.3: *Sample banks by county*

This figure presents the geographic distribution of sample banks in the US as of 2002Q1. The map shows the counties which have banks that are fully concentrated in the county, along with number of the concentrated banks operating in these counties.

information on the income of residents living in that particularly county. The income numbers are based on county income data on the addresses reported on individual income tax returns filed with the IRS. I include data on wages from the Quarterly Census of Employment and Wages (QCEW) that tracks wage and employment statistics for individuals who work, but who do not necessarily live in that county. I include data on per capita income from the Bureau of Economic Analysis (BEA). Data on the change in the number of private establishments to capture performance of business in the county come from the QCEW. To capture shocks in the agriculture sector, I use data on farming income at the county level from the BEA.

I obtain data on business/non business bankruptcies from the Administrative Office of the U.S. Courts. This source gives statistics on the number of bankruptcies filed at a county level. Finally, I obtain from the Census Bureau data on the demographics of various variables at the county level: of population, senior share, non-white share, the vacancy rate in housing, poverty levels of households, and the fraction of the population that has less than a high school education.

1.4.3 Bank Data and Summary Statistics

For the banks in my sample, I collect quarterly balance sheet information from the Bank's Quarterly Reports of Condition and Income (the "call reports"). Data on loan performance are the loans that are thirty days, ninety days or more past due and those that fall under non-accrual, all scaled by the outstanding loan balance. The size of the bank is measured as the natural logarithm of the bank's assets. As suggested by Ryan (2012), data on percentage of loans in real estate (RE) and under commercial and industrial (CI) is collected to reflect variation in bank loan portfolio compositions. I further collect data on loan to asset ratio along with securities to asset ratio, where securities are the book value of all held to mature securities.

I use other bank-level variables that are used in prior research to predict future loan losses. I use interest receivables for income accrued but not yet collected on loans (Prescott and Walter, 2015). Interest receivable also capture the difference between the sum of loans outstanding and the present value of contractually promised future payments on loans (Ryan, 2007). I capture loan yields on the loan portfolio, which should reflect the risk of default and the loss in the underlying risks in the portfolio. I define this measure as the ratio of tax-equivalent interest income divided by total loans at the end of the quarter following Jiang *et al.* (2016).

Table 1.2 reports summary statistics for the sample of bank-quarter-county in the period 1996 – 2012. The mean sample bank has \$104 million in total assets, while the median has \$64 million in assets. I use the logarithm of total assets in the models I run. The median and mean of total loans are \$38 million and \$66 million respectively. The mean rolling four-quarter net-chargeoffs (NCO) is $\$269,000$, while the mean rolling four-quarter NCO + non-performing loans (NPL) at the end of the four quarters is \$1.4 million. The mean and the median loan loss allowance of the concentrated banks are $\$892,000$ and $\$483,000$ respectively.

The mean sample bank has about 0.4% in loans that are ninety day past due, and 1% in loans that are in non-accrual. The mean loan to asset percentage is 61%, while the median is 62% and shows the fact that the loans are a large percentage of these banks' assets. The mean percentage of real estate loans is 64%, while the mean percentage of commercial and industrial loans in the sample is 14%.

I measure all county variables in terms of quarterly changes. The mean housing price growth for the sample counties is 1.1%, and the mean unemployment rate in the sample is 5.5%. The change in farm earnings is a decrease of about 4.2%, while the mean wage growth in the county from the IRS 1.1%.[24] The mean and median of the ratio of all business bankruptcies to the establishments in the county are 62% and 44% respectively. The mean home price growth for the counties in the sample is 1.1%. The change in farm earnings is a decrease of about 4.2%.

1.5 Empirical Results

This section provides details of the prediction algorithm and the loss function used to measure the accuracy of my allowance models (section 1.5.1), tests of the limited information model (section 1.5.2), full information model (section 1.5.3), magnitude of impact from using these models (section Section 1.5.4), and a discussion of the results from these models (section 1.5.5).

[24]The IRS data are available only on annual basis. I follow Mian and Sufi (2011) in interpolating the IRS data in obtaining the quarterly data.

Table 1.2: *Summary Statistics*

Statistic	N	Mean	P25	Median	P75	St. Dev.
Bank-level data						
Assets	306,666	103,917.90	35,168.2	64,531	120,032	150,108.80
Total Loans	306,666	65,556.07	19,209	38,212	75,499	99,042.87
Pct Four-qtr NCOs + NPL	306,666	1.83	0.31	0.98	2.26	10.76
Four-qtr NCOs + NPL	306,666	1,359.48	92	359	1,038	10,795.08
Chargeoffs	306,666	157.40	1	20	94	761.38
Recoveries	306,666	30.53	0	5	24	146.21
Allowance	306,666	892.17	242	483	974	1,543.63
Pct Allowance	306,666	1.49	0.99	1.28	1.73	1.14
30 days PD	306,666	0.97	0.00	0.23	1.41	1.57
90 days PD	306,666	0.42	0.00	0.05	0.41	1.04
Nonaccrual	306,666	1.02	0.00	0.33	1.19	1.96
Delta 30 days PD	306,666	0.03	−0.08	0.00	0.13	1.29
Delta 90 days PD	306,666	0.00	−0.05	0.00	0.06	0.81
Delta Nonaccrual	306,666	0.02	−0.08	0.00	0.06	0.93
Pct RE Loans	306,666	64.47	49.36	65.98	81.24	21.95
Pct CI Loans	306,666	13.97	6.36	12.01	19.13	10.97
Loans to Assets	306,666	60.67	50.69	62.12	72.33	16.12
Securities to Assets	306,666	25.67	13.63	24.04	35.67	15.88
Interest Receivables	306,666	0.71	0.40	0.57	0.89	0.45
Loan Yields	306,666	7.95	6.75	7.95	9.06	1.63
County-level data						
County Unemployment	306,506	5.54	3.80	5.00	6.70	2.57
HPI	306,487	129.26	106.29	122.95	149.23	33.06
HPI Growth	306,487	0.01	0.00	0.01	0.02	0.02
Establishments	306,666	0.00	−0.00	0.00	0.01	0.02
Wage Growth	305,896	0.01	0.00	0.01	0.02	0.02
Farm Earnings	302,229	−4.24	−7.00	−0.11	8.67	1,396.92
Per Capita	305,882	1.01	0.46	0.99	1.50	1.20
Bus Bankr	306,666	0.62	0.24	0.44	0.78	0.68

This table presents summary statistics for sample bank-level and county-level variables. The variables are in the period between 1996 and 2012. These variables are used in the limited information model and full information model. The data sources are identified in the appendix Table A.1.

I conclude this section by discussing a setting where the accounting rules would have been particularly helpful in states that suffered from severe banking crisis (section 1.5.6).

1.5.1 Predicting Allowance from Lasso

I develop a new strategy using lasso to obtain counterfactual estimates of allowances. The sample used in the study is a panel of bank-quarter observations in the periods 1996 – 2012. I estimate the lasso in a rolling-window estimation period, and then predict allowances in a test period which are one-quarter and two-quarter ahead out-of-sample.[25] The dependent variable in the lasso is future loan losses, as measured by total four quarter net chargeoffs + loans ninety-days past due + non-accrual. The variables used for predictions in the limited information and full information models are presented in Table 1.3. Using the lasso for estimating allowance under the limited and full information set allows me to dynamically create proxies for credit risk at each period using bank and county level economic information.

Table 1.3: *Variables used in the Models*

LINM	FINM
30 days PD	30 days PD
90 days PD	90 days PD
Nonaccrual PD	Nonaccrual PD
30 days PD (t-3)	30 days PD (t-3)
90 days PD (t-3)	90 days PD (t-3)
Nonaccrual PD (t-3)	Nonaccrual PD (t-3)
Loans to Assets	Loans to Assets
Securities to Assets	Securities to Assets
Pct RE Loans	Pct RE Loans
Pct CI Loans	Pct CI Loans
Log Assets	Log Assets
NCO_t	NCO_t
NCO_{t-3}	NCO_{t-3}
NCO_{t-6}	NCO_{t-6}
NCO_{t-9}	NCO_{t-9}
	Loan Yields
	Interest Receivables

(*Continued*)

[25] For example, the allowance estimates in 2007Q1, the estimation period ends at 2006Q4, while the one-quarter out-of-sample test period is 2007Q1.

Table 1.3 – *Continued*

LINM	FINM
	Unemp Rate$_t$
	Unemp (t-3)
	Unemp (3-6)
	Unemp (6-9)
	Unemp (9-12)
	HPI (t-3)
	HPI (3-6)
	HPI (6-9)
	HPI (9-12)
	Establishments (t-3)
	Establishments (3-6)
	Establishments (6-9)
	Establishments (9-12)
	Wage growth (t-3)
	Wage growth (3-6)
	Wage growth (6-9)
	Wage growth (9-12)
	Farm Earnings (t-3)
	Farm Earnings (3-6)
	Farm Earnings (6-9)
	Farm Earnings (9-12)
	Per Capita (t-3)
	Per Capita (3-6)
	Per Capita (6-9)
	Per Capita (9-12)
	Bus Bankr (t-3)
	Bus Bankr (3-6)
	Bus Bankr (6-9)
	Bus Bankr (9-12)

This table reports the variables that are used for the prediction models developed in the paper. The LINM column shows variables used in the limited information model, and the FINM column shows variables used in the full information model. The subscript variables refer to lags, while the parenthesis show the difference in variable between the time period specified. For example, NCO_{t-3} is the one-quarter lag in the NCO, while HPI (t-3) is the three quarter housing price growth between the periods t and $t-3$.

The lasso performs, what is termed variable selection, by shrinking the coefficients of some of the predictor variables to zero. I assume that this is how the allowance estimation would work in practice, where bank managers would weigh only a subset of the information that is relevant to predicting future loan losses.[26]

[26]For example, when using the full information set, a bank that operates in a county that sees house price declines would weigh that factor more heavily, compared to a bank that operates in a county that has high bankruptcies but a relatively stable housing market.

As argued in Section 1.3.4, in implementing the lasso, it is critical to select a suitable value of λ. At each iteration of the estimation, I perform a ten-fold cross validation using the relevant predictor and loan loss measure. The cross validation method is commonly used in the ML literature (see Hastie *et al.*, 2009).[27] The algorithm splits the historical data into n-subsets (or folds). For a set of λ, the algorithm is estimated on the $n-1$ folds and it then validates which value of λ produces the best prediction in the n-th fold. This process is repeated until the model is tested on each of the n folds, which produces the cross-validated λ. In order for my model to have reasonable estimation window, I start my first predictions in 2002Q1. Therefore, in the first iteration, the model looks at all data from 1996Q1 to 2002Q1 in performing the estimates. I then make predictions using the coefficients from the lasso one quarter and two quarters ahead. In the next iteration, I increase the window by a quarter, and include 2002Q2.

In evaluating the prediction accuracy of the lasso models and the ILM, the main loss function I use is as proposed in machine learning, and more broadly in the literature on forecasting (see Hamilton, 1994, chap. 4). I use the out-of-sample predictions for the testing, and use the root mean square error (RMSE), which is the square root of the mean squared error.[28] This measure is the difference between the actual loan losses (four quarter net-chargeoffs plus non-performing loans) for a bank in a quarter, and the corresponding estimated allowances in that quarter. I calculate the measure for each bank in each period. Then I square the difference, and the mean of this loss measure is the RMSE. I repeat this procedure for the allowance under ILM, for each bank in each period. The model with the lower RMSE predicts subsequent net-chargeoffs more accurately, and is regarded as a better model.

[27]For my lasso estimates, I use the `glmnet` package in R that implements a class of regularization algorithms.

[28]The mean squared error is $\mathbb{E}(Y_{t+1}-\widehat{Y}_{t+1}|t)^2$ based on the estimated model at t. Here Y is the actual net-chargeoff rate, while the value $\widehat{Y}_{t+1}$ is the model predicted allowance as a percentage of loans. For calculating the MSE under the incurred loss model, I consider a prediction model that takes loan loss allowance as the only independent variable. I use the predicted value from this estimated model as the allowance under ILM.

1.5.2 Evaluation of the Limited Information Model

I present the results in a series of tables that report the RMSE of the lasso-estimated limited information model and the ILM. If the limited information model outperforms the ILM, I expect to observe its RMSE to be lower than the ILM.

I estimate the models on quarterly bank data, with predictions of allowances beginning in 2002. I present evidence by aggregating the quarterly RMSE at the yearly level. I report the cumulative RMSE, and the yearly RMSE. The cumulative RMSE for year t in the table is cumulative means of RMSE for all years less than or equal to t. The yearly RMSE for t is the means of the RMSE for year t.

The results for the performance evaluation between the limited information model allowance (LINM) and the ILM allowance are reported in Table 1.4 and Table 1.5. The key result is that allowance from the limited information model predicts future loan losses more accurately, and outperforms the allowance under the ILM. The RMSE estimates in both one-quarter and two-quarters out-of-sample model estimates using lasso are less than those of the ILM.

The cumulative RMSE for the one-quarter ahead and two-quarter ahead predictions are in panel A and panel B of Table 1.4, respectively. The RMSE for the limited information model (LINM) for the overall sample between 2002 and 2012 is 3.22 for one-quarter ahead, and 3.26 for two-quarters ahead. The RMSE for the ILM allowance in this period is 6.07 for one-quarter ahead, and 6.66 for two-quarters ahead. The cumulative RMSE for the one-quarter ahead limited information model estimated for the sample ending in 2009 is 3.65, compared to the ILM which is 6.74.

The results in Table 1.5 show the one-quarter and two-quarter RMSE for estimates for the individual years. In contrast to Table 1.4, the RMSE in Table 1.5 are the aggregated values by year for the RMSE values. Overall, the results consistently show that the limited information model has better accuracy than the ILM in predicting future loan losses. The performance of both the limited information model and the ILM suffer in 2008 during the mortgage crisis, but the limited information model continues to outperform.

The evidence reported indicates with that it is possible to construct a prediction model that outperforms the ILM without having to expand the information set. I interpret the performance difference attributed to using the lasso model, which relaxes any restrictions on how the information set can be used while estimating the allowances. The lasso model efficiently weighs the limited information set to improve the overall prediction accuracy. The implication is that an objective and simple approach using the lasso to estimate allowance brings performance gains to the ILM, without having to expand the information set.

1.5.3 Evaluation of the Full Information Model

The results of the performance evaluation of the full information model (FINM) and the limited information model (LINM) are reported in Table 1.4 and Table 1.5. The models are estimated on a quarterly bank-county economic data. I use the limited information model as the baseline to evaluate the full information model's performance. If incorporating broader information in the LINM lasso model brings significant benefits in allowance estimation, I expect to observe the RMSE of the full information model significantly differ from RMSE of the limited information model.

However, I find that the performance of the limited information model does not change significantly on expanding the information in the full information set. The RMSE for both the one-quarter and two-quarter out of samples do not exhibit significant differences. Nevertheless, the full information model (FINM) does outperform the ILM. The cumulative RMSE for the one-quarter ahead and two-quarters ahead predictions are in panel A and panel B of Table 1.4, respectively. The RMSE for the full information model for the overall sample between 2002 and 2012 is 3.19 for one-quarter ahead, and 3.24 for two-quarters ahead. The RMSE for the limited information model is 3.22 for one-quarter ahead and 3.26 for two-quarters ahead.

The yearly estimates exhibit similar pattern as shown in Table 1.5. Panel A presents the one-quarter ahead RMSE, while panel B presents the two-quarters RMSE for the estimates aggregated by the year. The one-quarter RMSE for the full information model in 2007 is 2.2, while the limited information model is 2.304. The magnitudes of these differences are similar in 2012 with the RMSE for the full information model being 1.85 and for the limited information model 1.89. Consistently, I find that the broader information does not does not bring performance gains to the

Table 1.4: *Allowance Model Cumulative Performance By Year*

Panel A: One Quarter Ahead											
Year	2002	2003	2004	2005	2006	2007	2008	2009	2010	2011	2012
FINM	1.703	1.521	1.427	1.384	1.387	1.522	2.736	3.635	3.482	3.333	3.199
LINM	1.558	1.440	1.373	1.348	1.362	1.519	2.752	3.656	3.503	3.354	3.221
ILM	1.755	1.636	1.574	1.553	1.581	1.788	5.238	6.743	6.492	6.267	6.068
Panel B: Two Quarters Ahead											
Year	2002	2003	2004	2005	2006	2007	2008	2009	2010	2011	2012
FINM	1.655	1.490	1.398	1.368	1.398	1.593	3.331	3.673	3.513	3.352	3.245
LINM	1.502	1.412	1.348	1.333	1.377	1.597	3.346	3.692	3.533	3.371	3.266
ILM	1.761	1.656	1.597	1.589	1.649	1.964	6.843	7.373	7.085	6.826	6.656

This table presents evidence on the performance of the model predicted allowance relative to the current GAAP's ILM in their accuracy in predicting future losses. It shows the cumulative root mean squared error by year. LINM is the root mean squared error of the predicted allowance from the lasso model that takes the limited information as input. FINM is the root mean squared error of the predicted allowance from the lasso model that takes the full information as the input. ILM is the root mean squared error of allowance under current GAAP from the financial statements. The error are generated from rolling-window predictions and measured as the accuracy of the allowance estimates to predict future net-chargeoffs. Panel A examines the prediction allowance using one-quarter out of sample. Panel B examines the predicted allowances using two-quarter out of sample.

Table 1.5: *Allowance Model Performance By Year*

Panel A: One Quarter Ahead											
Year	2002	2003	2004	2005	2006	2007	2008	2009	2010	2011	2012
FINM	1.703	1.339	1.238	1.256	1.398	2.200	10.019	9.923	2.264	1.993	1.851
LINM	1.558	1.321	1.240	1.272	1.418	2.304	10.152	9.980	2.280	2.009	1.893
ILM	1.745	1.499	1.431	1.482	1.703	2.653	11.597	11.405	3.188	2.675	2.350
Panel B: Two Quarters Ahead											
Year	2002	2003	2004	2005	2006	2007	2008	2009	2010	2011	2012
FINM	1.655	1.325	1.214	1.276	1.519	2.567	13.762	6.063	2.235	1.899	1.818
LINM	1.502	1.322	1.220	1.288	1.551	2.697	13.843	6.114	2.254	1.913	1.867
ILM	1.747	1.538	1.459	1.564	1.891	3.190	16.868	7.854	3.203	2.578	2.364

This table presents evidence on the performance of the model predicted allowance relative to the current GAAP's ILM in their accuracy in predicting future losses. It shows the root mean squared error by year. LINM is the root mean squared error of the predicted allowance from the lasso model that takes the limited information as input. FINM is the root mean squared error of the predicted allowance from the lasso model that takes the full information as the input. ILM is the root mean squared error of allowance under current GAAP from the financial statements. The error are generated from rolling-window predictions and measured as the accuracy of the allowance estimates to predict future net-chargeoffs. Panel A examines the prediction allowance using one-quarter out of sample. Panel B examines the predicted allowances using two-quarter out of sample.

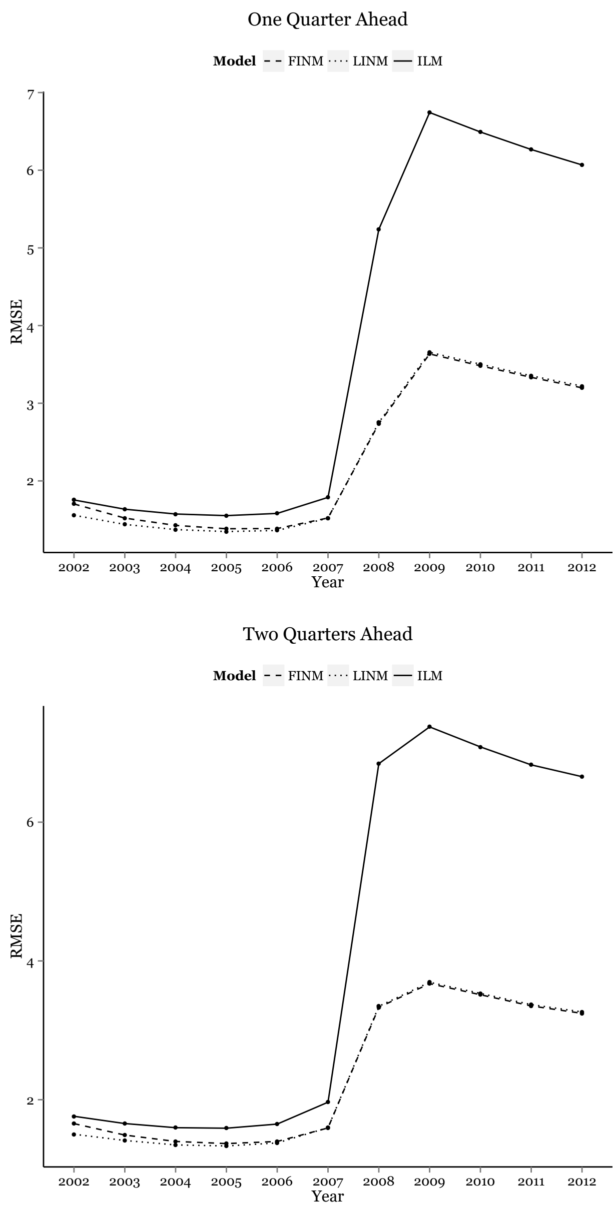


Figure 1.4: *Allowance Model Cumulative Performance By Year.*

This figure presents evidence on the performance of the model predicted allowance relative to the current GAAP's ILM in their accuracy in predicting future losses. It shows the cumulative root mean squared error by year as reported in Table 1.4. LINM is the root mean squared error of the predicted allowance from the lasso model that takes the limited information as input. FINM is the root mean squared error of the predicted allowance from the lasso model that takes the full information as the input. ILM is the root mean squared error of allowance under current GAAP from the financial statements.

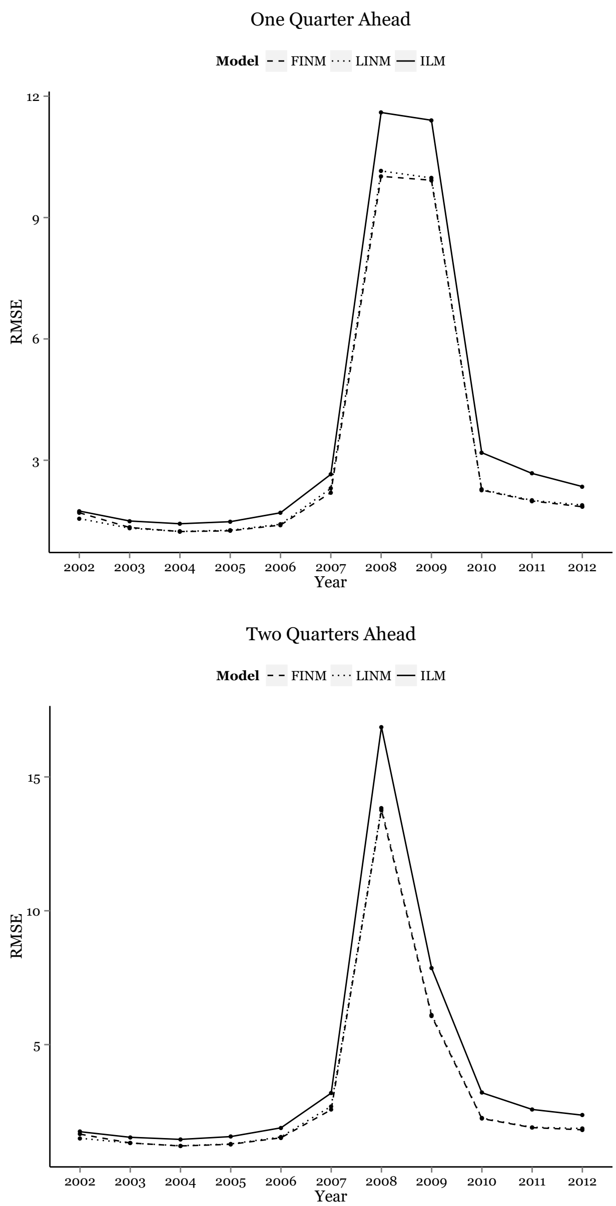


Figure 1.5: *Allowance Model Performance By Year*

This figure presents evidence on the performance of the model predicted allowance relative to the current GAAP's ILM in their accuracy in predicting future losses. It shows the root mean squared error by year as reported in Table 1.5. LINM is the root mean squared error of the predicted allowance from the lasso model that takes the limited information as input. FINM is the root mean squared error of the predicted allowance from the lasso model that takes the full information as the input. ILM is the root mean squared error of allowance under current GAAP from the financial statements.

full information model compared to the limited information model.

In summary, the evidence reported is consistent with expanding information set in estimating allowances to the limited information model, as suggested in the CECL, provides no significant benefit in model performance relative to the limited information lasso model. I attribute the result to the lasso approach that flexibly weighs the limited information set, as was available in the ILM, to produce an accuracy similar to that obtained using the full information set.

1.5.4 How much more Allowance?

My statistical tests of the allowance models thus-far focused on the accuracy of the estimates to predict future loan losses. I now discuss the impact of the estimated allowance by measuring the magnitude of the difference between the actual loan loss allowances that were accrued by banks, and my model predicted allowance. I measure the difference as a percentage of loans, and compare the dollar value of this measure relative to the actual allowance accrued reported in the bank financial statements. I calculate this measure for each sample bank in each quarter, and aggregate it to the year. The idea is that if managers had used the limited information model or full information model to estimate allowances, they would have increased or decreased the allowance recognition they had using the ILM.

Figure 1.6 summarizes the results for the limited information model for the one-quarter (top panel) and two-quarters ahead (bottom panel) predictions. The difference in the allowances as percentage of loans in 2008 is 0.4%, and in 2009 is 0.8%. The mean bank in the sample accrues 1.5% of loans as allowance as presented in Table 1.2. I find that using the limited information model one-quarter ahead estimates in the periods between 2002 – 2012, the mean bank in the sample would have increased its allowance recognition by 22%, relative to what it had. This impact is economically meaningful.

In the period between 2002 – 2003, the limited information model predicts that banks would have increased allowances relative to ILM. Then for the periods from 2004 to 2005, this difference is negative, suggesting the banks would have lowered allowance using the limited information model relative to the ILM. But entering the crisis, beginning in 2006, the limited information

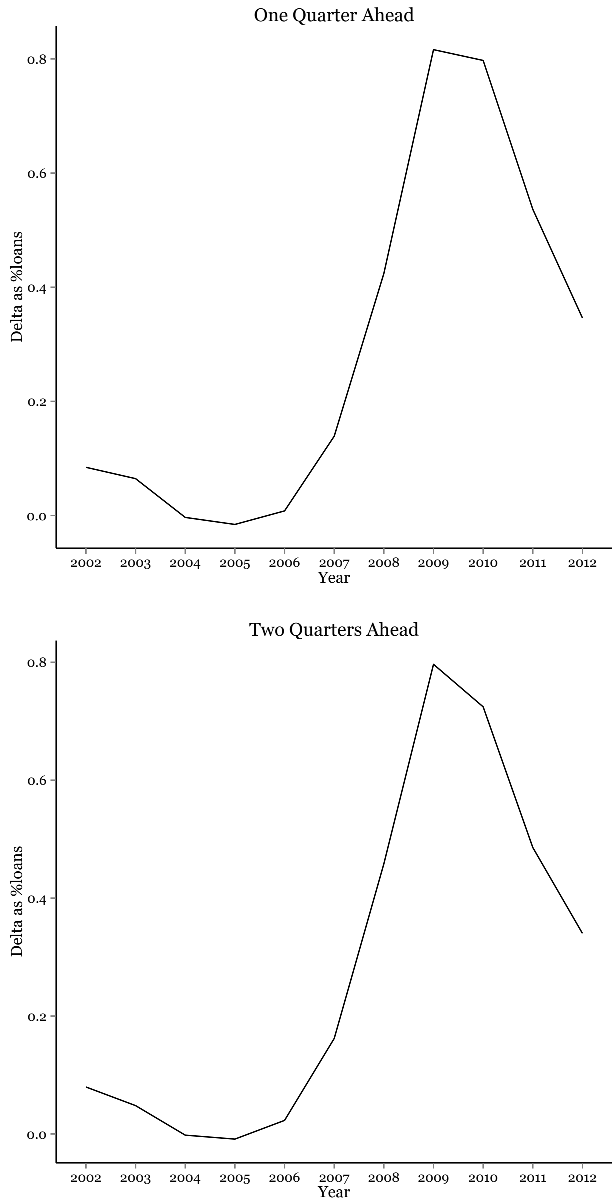


Figure 1.6: *Difference between Predicted Allowance and Actual allowance*

This figure presents magnitude difference between the allowance estimated from the LINM and actual loan loss allowance from the bank's financial statements. The difference is calculated for each bank in the sample for each quarter, and mean difference is aggregated by year. The mean difference per year is plotted as a percentage of outstanding loans. LINM is the allowance model estimated from the lasso model that takes the limited information as the input.

model would have predicted higher allowances than what the banks had under the ILM. The limited information model predicts banks should have increased their allowance by 10% relative to what they had under ILM in 2007. I find similar results for the two-quarter ahead limited information model predictions.

In analysis not presented, I find similar economically meaningful results for allowances predicted by the full information model. On average, banks would have increased their allowance by 22% for the period between 2002 – 2012. However, in contrast to the limited information model, the full information model would have predicted lower allowances than that of the ILM during the early part of the analysis until about 2005. But similar to the lower information model, banks under the full information model would have increased their allowance entering the crisis beginning in 2006. Using the full information model, banks should have increased their allowance by 23% in 2007.

The predicted allowances from the limited information and full information model together suggest that using the ILM for allowances understated the value of sample banks' portfolios as presented in Figure 1.6. Moreover, this inaccurate level of allowance under the ILM has a direct impact on bank capital. The difference in the allowance estimates between the models presented in the figures reveal the unclear picture of how capital adequate the banks really were. These aspects suggest that with the shortcomings of the ILM allowance model, banks and its supervisors would have had a false sense of the risks in the bank balance sheets. As a result, banks could have continued to lend and, hold excessively risky loan positions on their balance sheets. Besides, the allowance estimates from the limited and full information models have the ability to provide early warning about changing economic conditions as seen in the increase in allowances beginning in 2006 for the financial crisis.

1.5.5 Discussion of the Limited and Full Information Models

This section explores the implications of my empirical results from the limited and full information models towards improving the current accounting rules. In particular, I consider two aspects of the ILM and the CECL – the judgment allowed under the rule, and the information that can be used in allowance estimation.

The evidence that limited information model outperforms ILM raises the possibility that the managers were reluctant to use the information they had access to under the ILM. Their judgment seems to have not been exercised in an unbiased way. The limited information model implements a loan loss accounting rule without any constraints on how the information is used; but removes the judgment on the part of the managers relative to the ILM. Given that the limited information model predicts future loan losses more accurately, it raises the question of whether affording managers with more discretion in my setting, improves the quality of accounting and enhance transparency. My findings document the effect from the discretion.

The CECL rule stresses using broader information to estimate allowances. From the results of the full information model, it is not clear that broader information alone will improve the performance. As a result, more guidance would be helpful for managers on how to implement the accounting rule. The CECL, as issued, does not specify a single method for measuring the losses; and leaves it to the banks to develop methods that help in objectively applying the principles of the CECL.

In the analyses presented in the paper using the limited information and full information model, I show how to implement a CECL rule, and more broadly an allowance rule, using the lasso approach. The lasso offers a simple, and objective approach that clearly outperforms current GAAP. By incorporating a broader set of county economic data, I show how a rule can be implemented, that is also verifiable and can be consistently applied over time. Furthermore, this approach can be easily implemented in practice. These findings will be of interest to standard setters, bank managers, and auditors.

1.5.6 Allowances in States that Suffered Severe Recession

In this section, I consider banks in states that suffered severe recession with many bank failures. I then examine the performance of the LINM, FINM, and ILM for these banks. The idea is to study how the accounting information in allowances could have played a larger role and provided early warning of the changing economic conditions. Hence, the allowance estimates from the models

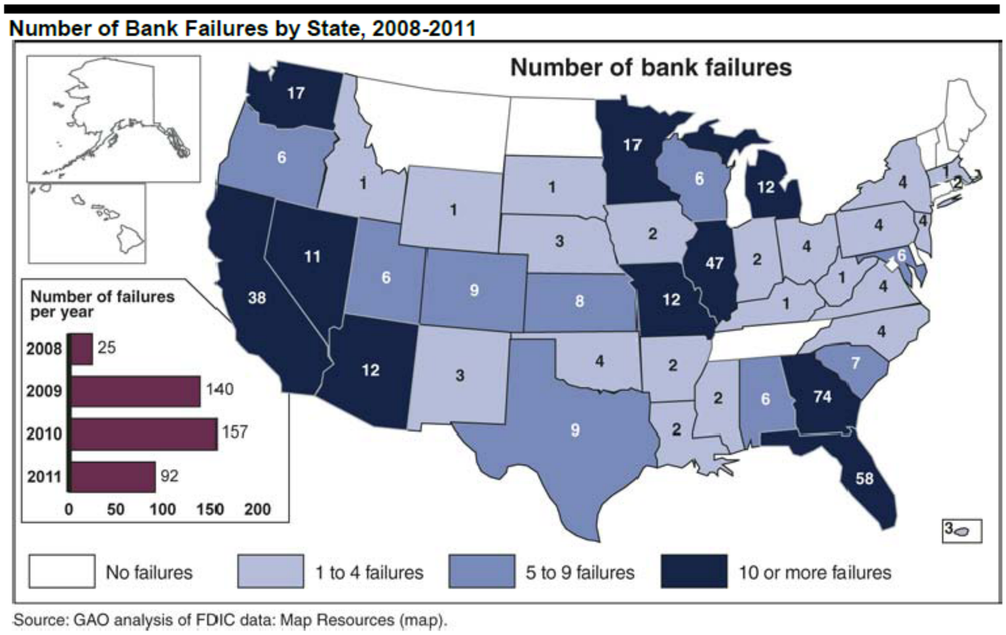


Figure 1.7: *Distribution of Failed Banks in the US.*

This figure shows the number of failed banks by state between 2008 – 2011.

would possibly acted as a trigger and allowed bank supervisors to heed to the warning, and act on the risk accumulation in these banks thus mitigating the threat to financial stability.

For this analysis, I first identify states that suffered severe banking failures during the mortgage crisis. I use the list of states from Figure 1.7, which shows the distribution of failed banks in the US. I consider states with at least 10 bank failures in the period 2008 – 2011. The states that fall in this category are Arizona, California, Florida, Georgia, Illinois, Michigan, Minnesota, Missouri, Nevada, and Washington. For these severe crises states, I expect to observe larger effects in the performance difference between the ILM relative to the limited information model, and the full information model. I also hypothesize that the magnitude difference between the estimated allowances and the actual reported loan loss allowances should be higher than the ones documented in the previous case for the sample of all banks. Note, the classification of the states happen based on ex post performance of the banks.

I estimate the allowances using the limited information model and the full information model for only the subsample of banks that operate in the identified states. I compare the RMSE for the models relative to the ILM. The results are reported in Table 1.6. Panel A in this figure shows the RMSE for one-quarter ahead predictions, and panel B the two-quarters ahead predictions. The evidence in the figure suggests that the limited information model continues to outperform the ILM, and that adding new information does not bring significant benefits. However, the ILM severely underperforms thus widening the wedge between the performance of the ILM and my estimated models.

The one-quarter cumulative RMSE of the limited information model in the period 2002 – 2012 is 4.479, while the RMSE for the ILM is 10.007. The RMSE for the full information model in this period is 4.574. The RMSE in the periods 2002 – 2008 for the limited information model is 3.691, and for the ILM is 8.658. The RMSE by year for the models are reported in Table 1.7.

These prediction accuracy estimates translate to large magnitude differences in the allowance amounts estimated by my models and the ILM's allowances. Similar to the analysis in Section 1.5.4, I calculate the difference between the actual loan losses and my model predicted allowances for each bank in the subsample in each quarter. I then aggregate the difference to the year. The findings are presented in Figure 1.10. I find the mean bank that operated in the states with severe banking crisis should have increased its allowances by 33% using the limited information model, relative to the actual allowance they had accrued in the period between 2002 – 2012. This increase is higher compared to the estimates based on a sample of all banks. Under the full information model, banks should have increased their loan loss recognition by 29%.

Furthermore, I find that using both the limited information model and full information model, banks would have started to increase their allowance beginning in 2006. For example, the limited information model predicts that the banks should have increased their allowance by 10% in 2007 entering the crisis. The full information model, on the other hand, would have predicted the banks to have increased by 35%. During the initial part of the estimation sample, there are no significant differences between the ILM allowance and the limited information model allowance. But in the same period, the allowance under the full information model would have suggested that the allowance be decreased.

Overall, the evidence presented for the subsample of banks that operated in states that suffered

Table 1.6: *Allowance Model Cumulative Performance by Year in US States with Severe Banking Crisis*

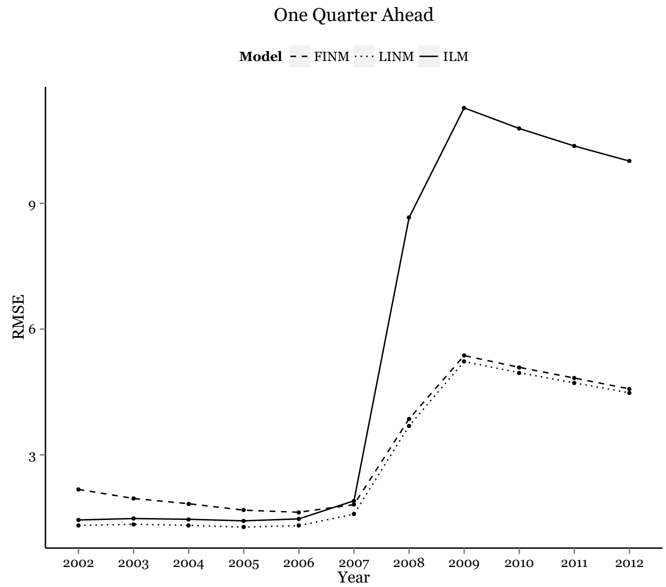
Panel A: One Quarter Ahead											
Year	2002	2003	2004	2005	2006	2007	2008	2009	2010	2011	2012
FINM	2.176	1.960	1.835	1.681	1.631	1.813	3.858	5.369	5.088	4.832	4.574
LINM	1.317	1.346	1.320	1.277	1.313	1.586	3.691	5.226	4.959	4.717	4.479
ILM	1.447	1.482	1.464	1.427	1.473	1.904	8.658	11.277	10.785	10.367	10.007
Panel B: Two Quarters Ahead											
Year	2002	2003	2004	2005	2006	2007	2008	2009	2010	2011	2012
FINM	2.229	2.001	1.858	1.695	1.683	1.947	4.911	5.453	5.168	4.877	4.673
LINM	1.282	1.330	1.310	1.269	1.357	1.716	4.728	5.292	5.025	4.750	4.563
ILM	1.471	1.520	1.507	1.470	1.583	2.201	11.509	12.378	11.827	11.354	11.048

This table presents evidence on the performance of the model predicted allowance relative to the current GAAP's ILM in their accuracy in predicting future losses in sample banks that operated in US states that suffered severe banking crisis. The states are identified by the ones that have more than 10 bank failures in the period 2008 – 2011. The table shows the cumulative root mean squared error by year. LINM is the root mean squared error of the predicted allowance from the lasso model that takes the limited information as input. FINM is the root mean squared error of the predicted allowance from the lasso model that takes the full information as the input. ILM is the root mean squared error of allowance under current GAAP from the financial statements. The error are generated from rolling- window predictions and measured as the accuracy of the allowance estimates to predict future net-chargeoffs. Panel A examines the prediction allowance using one-quarter out of sample. Panel B examines the predicted allowances using two-quarter out of sample.

Table 1.7: *Allowance Model Performance by Year in US States with Severe Banking Crisis*

Panel A: One Quarter Ahead											
Year	2002	2003	2004	2005	2006	2007	2008	2009	2010	2011	2012
FINM	2.176	1.743	1.586	1.218	1.434	2.721	16.130	15.941	2.840	2.530	2.001
LINM	1.317	1.374	1.269	1.149	1.454	2.951	16.326	15.970	2.825	2.540	2.097
ILM	1.446	1.510	1.421	1.295	1.650	3.364	18.772	18.321	3.842	3.179	2.577
Panel B: Two Quarters Ahead											
Year	2002	2003	2004	2005	2006	2007	2008	2009	2010	2011	2012
FINM	2.229	1.773	1.571	1.205	1.637	3.265	22.700	9.245	2.889	2.258	1.951
LINM	1.282	1.378	1.269	1.145	1.709	3.512	22.802	9.242	2.887	2.270	2.076
ILM	1.469	1.561	1.472	1.343	1.980	4.157	28.048	12.000	3.912	2.951	2.607

This table presents evidence on the performance of the model predicted allowance relative to the current GAAP's ILM in their accuracy in predicting future losses in sample banks that operated in US states that suffered severe banking crisis. The states are identified by the ones that have more than 10 bank failures in the period 2008 – 2011. The table shows the root mean squared error by year. LINM is the root mean squared error of the predicted allowance from the lasso model that takes the limited information as input. FINM is the root mean squared error of the predicted allowance from the lasso model that takes the full information as the input. ILM is the root mean squared error of allowance under current GAAP from the financial statements. The error are generated from rolling-window predictions and measured as the accuracy of the allowance estimates to predict future net-chargeoffs. Panel A examines the prediction allowance using one-quarter out of sample. Panel B examines the predicted allowances using two-quarter out of sample.



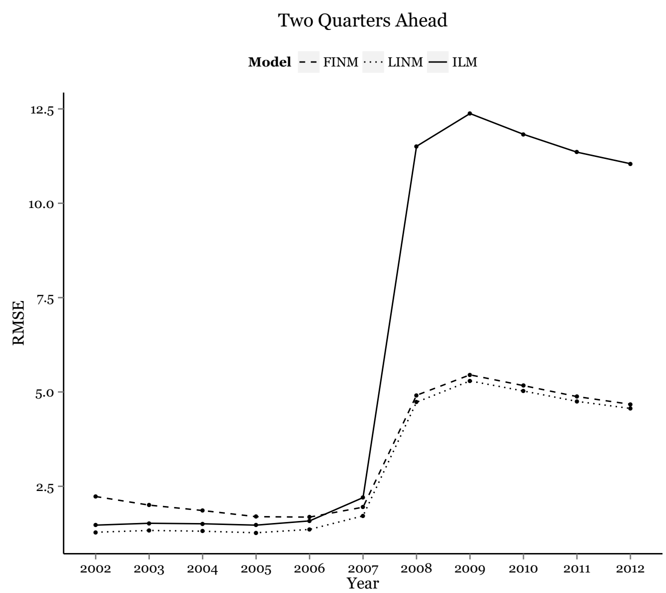


Figure 1.8: *Allowance Model Cumulative Performance By Year in US States with Severe Banking Crisis*

This figure presents evidence on the performance of the model predicted allowance relative to the current GAAP's ILM in their accuracy in predicting future losses in sample banks that operated in US states that suffered severe banking crisis. The states are identified by the ones that have more than 10 bank failures in the period 2008 – 2011. The figure shows the cumulative root mean squared error by year as reported in Table 1.6. LINM is the root mean squared error of the predicted allowance from the lasso model that takes the limited information as input. FINM is the root mean squared error of the predicted allowance from the lasso model that takes the full information as the input. ILM is the root mean squared error of allowance under current GAAP from the financial statements.

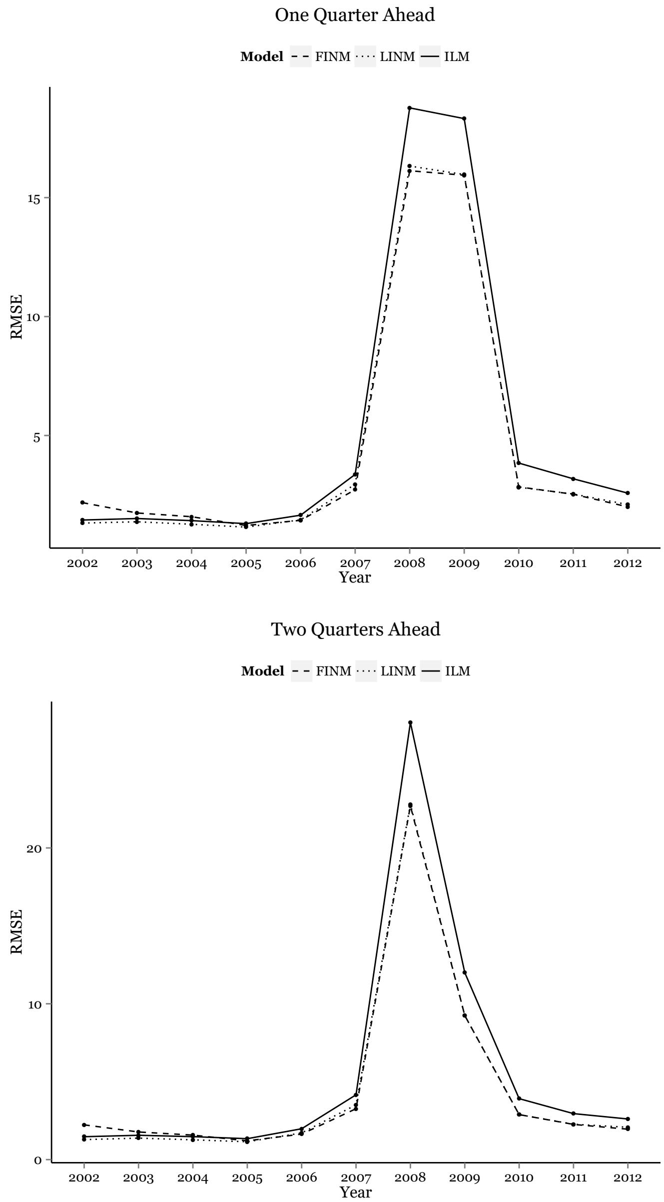


Figure 1.9: *Allowance Model Performance By Year in US States with Severe Banking Crisis*

This figure presents evidence on the performance of the model predicted allowance relative to the current GAAP's ILM in their accuracy in predicting future losses in sample banks that operated in US states that suffered severe banking crisis. The states are identified by the ones that have more than 10 bank failures in the period 2008 – 2011. The figure shows the root mean squared error by year as reported in Table 1.7. LINM is the root mean squared error of the predicted allowance from the lasso model that takes the limited information as input. FINM is the root mean squared error of the predicted allowance from the lasso model that takes the full information as the input. ILM is the root mean squared error of allowance under current GAAP from the financial statements.

severe crisis is consistent with that a limited information model would have more accurately predicted the loan losses than the ILM, and that full information model does not bring considerable performance gains. However, these allowance estimates could have possibly allowed the bank regulators and supervisors to act timely in limiting some of the bank failures.

1.5.7 How does the Limited Information Model Perform for Large Banks?

The analysis thus far focuses on a sample of county banks, which allows to map their the micro-economic information. But a natural extension is to examine the performance of the limited information model relative to the ILM for a broader sample of large banks. Not only do the large banks differ in terms of how they are regulated, but also tend to have different loan portfolio composition compared to the one-county banks. The large banks also have a capacity for greater risk and could estimate loan loss allowances using more sophisticated models. Therefore it is not clear ex ante how the limited information model would perform, given the performance of the ILM observed is a function of the models used in the estimation, and the incentives of the large banks.

To examine the limited information model for large banks, I consider banks that have more than $250 million in total assets (and operating in multiple counties).[29] I predict allowances for each sample bank in each quarter in the period 2002 – 2012, and compare the performance using the RMSE, as well as the magnitude difference between the estimated allowance and the actual loan loss allowance. The results from the out of sample RMSE are presented in Figure 1.11. I find the LINM continues to outperform the ILM and accurately predicts the loan losses. The RMSE for the LINM for the large banks for the overall sample between 2002 and 2012 is 2.12, while the RMSE for the ILM allowance in this period is 3.01. The cumulative RMSE for the one-quarter ahead (top panel) and two-quarter ahead predictions (bottom panel) are in Figure 1.12. The RMSE for the one-quarter ahead LINM estimated for the sample ending in 2009 is 2.03, compared to the ILM which is 3.04.

The magnitude differences between the LINM predicted allowance, and the actual allowance accrued by large banks is presented in Figure 1.13. The LINM predicts that the large banks would

[29] The mean bank in the sample has total loans of about $2.3 billion.

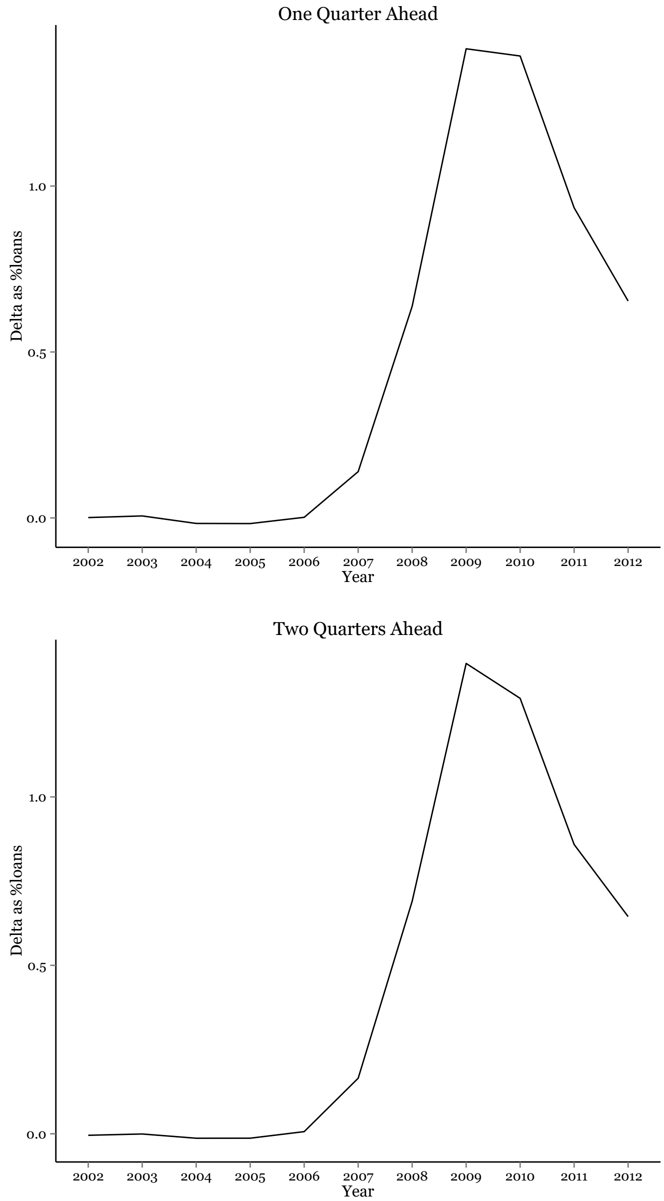


Figure 1.10: *Difference between Predicted Allowance and Actual allowance in US States with Severe Banking Crisis*

This figure presents magnitude difference between the allowance estimated from the LINM and actual loan loss allowance from the bank's financial statements. The figure shows the difference in sample banks that operated in US states that suffered severe banking crisis. The states are identified by the ones that have more than 10 bank failures in the period 2008 – 2011. The difference is calculated for each bank in the sample for each quarter, and mean difference is aggregated by year. The mean difference per year is plotted as a percentage of outstanding loans. LINM is the allowance model estimated from the lasso model that takes the limited information as the input.

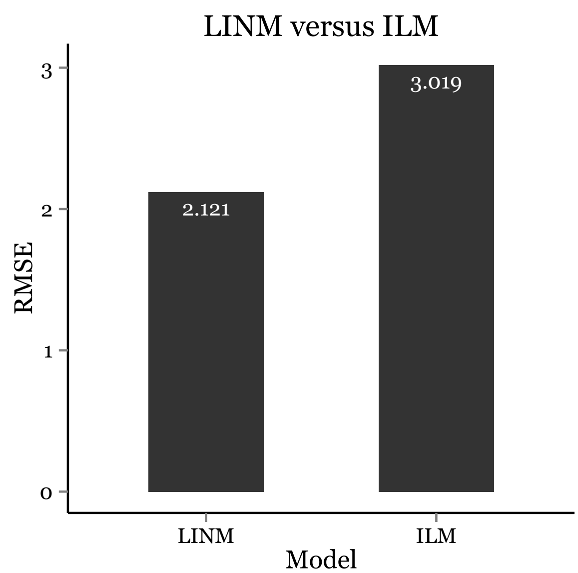


Figure 1.11: *Summary of Allowance Model Performance for Large Banks*

This figure presents evidence on the performance of the model predicted allowance relative to the current GAAP's ILM in their accuracy in predicting future losses for sample of large banks. The banks are identified with total assets greater than $250 million. It shows the cumulative root mean squared error for the model estimated and tested out of sample in the period 2002–2012. LINM is the root mean squared error of the predicted allowance from the lasso model that takes the limited information as input. ILM is the root mean squared error of allowance under current GAAP from the financial statements. Refer Figure 1.4 for further discussion on the models.

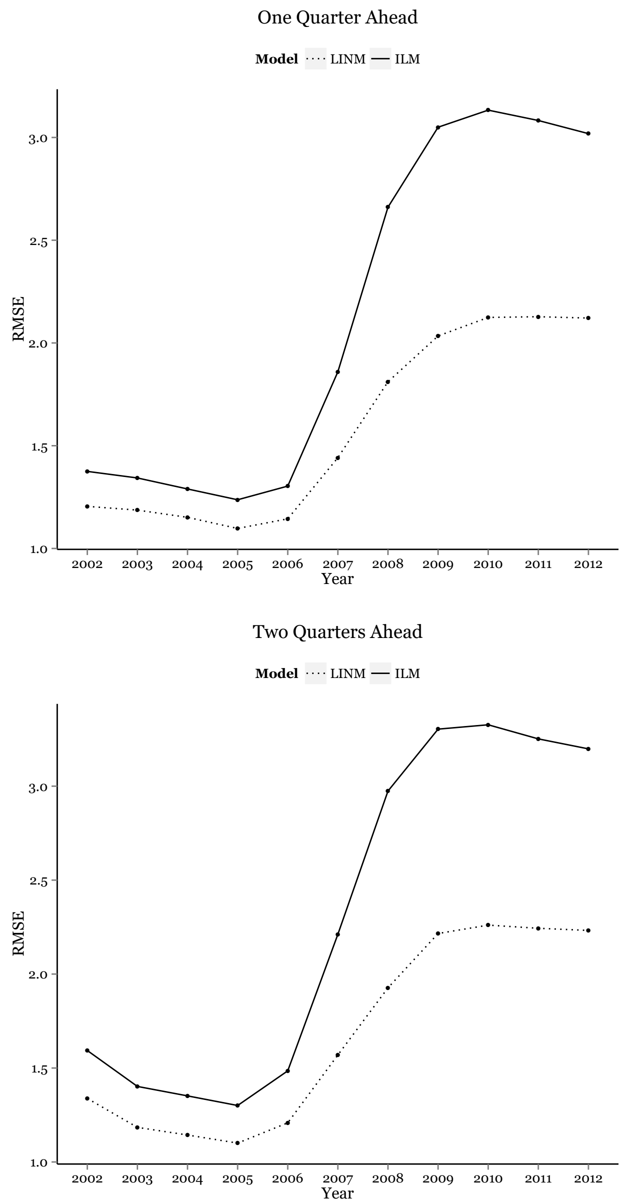


Figure 1.12: *Allowance Model Cumulative Performance By Year for Large Banks*

This figure presents evidence on the performance of the model predicted allowance relative to the current GAAP's ILM in their accuracy in predicting future losses for a sample of large banks. The banks are identified with total assets greater than $250 million. It shows the cumulative root mean squared error by year. LINM is the root mean squared error of the predicted allowance from the lasso model that takes the limited information as input. ILM is the root mean squared error of allowance under current GAAP from the financial statements.

have decreased their allowance in 2005, relative to what they had actually recognized. But early in the crisis, the LINM predicts that the large banks increase their allowance. The difference between the allowances in 2007 is 0.11% and in 2008 is .64%. These differences are economically meaningful. With the LINM one-quarter ahead estimates in the period between 2002 – 2012, the mean bank in the sample would have increased its allowance by 24% relative to what it had. The results are similar for the two-quarter ahead predictions.

The above analysis suggests that LINM continues to outperform the ILM for banks with large loan portfolios. This is consistent with the earlier findings using the one-county bank sample, adding further validity of the performance gains from using the LINM model.

1.5.8 How Robust are the Horizon of Future Losses?

The analysis in the paper predicts allowance based on four-quarter rolling net charge-offs plus non-performing loans. This measure is the dependent variable used in the lasso approach to capture bank managers' estimate of future losses. In reality, the actual horizon the banks use in estimating the losses are not observable. To test the robustness of my assumption, I estimate the models by extending the loss horizon in the prediction model to two-years and three-years. I then repeat the analysis and evaluate the out of sample performance accuracy of the limited information model and full information models, relative to the performance of the ILM in predicting the losses. I find that the results are qualitatively the same, and the limited information model continues to outperform the ILM. And the difference between the full information model does not bring any significant benefits. One advantage the four-quarter rolling window offers is that it allows to estimate allowances based on loans the bank already has underwritten as of the financial statements date.

1.6 What Drives the Performance Difference between LINM and ILM?

A principal finding in this paper is that the LINM outperforms the ILM. In this section I offer three explanations for the performance difference between the two models by 1) examining the weights assigned to the variables in the models (section 1.6.1); 2) determining where the effects from

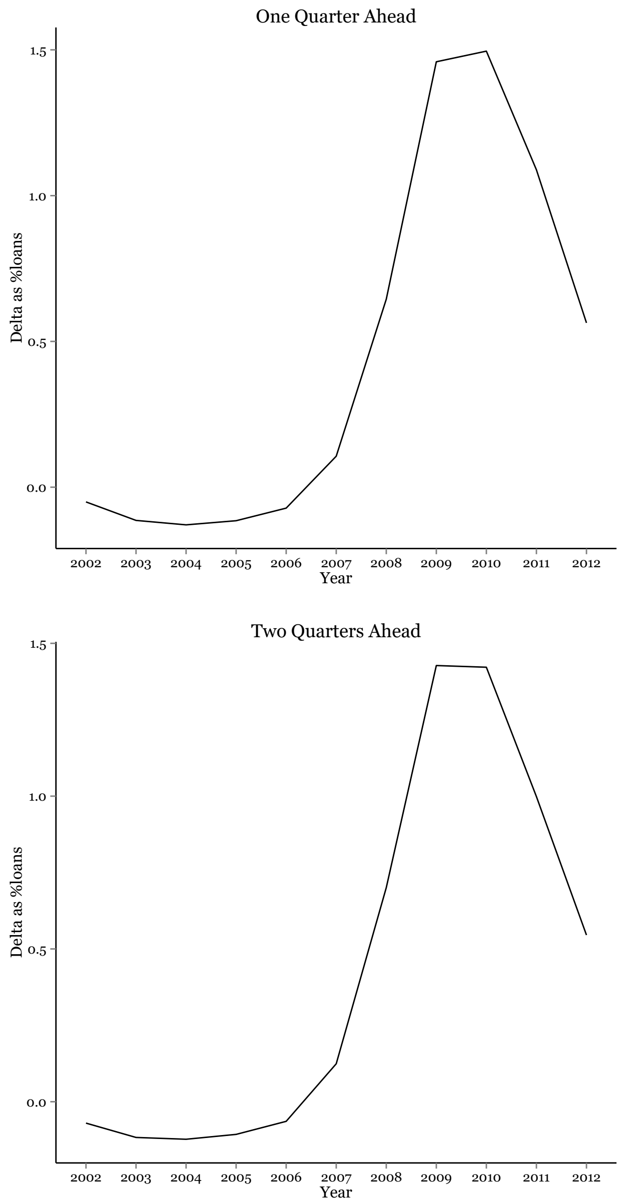


Figure 1.13: *Difference between Predicted Allowance and Actual allowance for Large Banks*

This figure presents magnitude difference between the allowance estimated from the LINM and actual loan loss allowance from the bank's financial statements. The figure shows the difference in sample of large banks. The banks are identified with total assets greater than $250 million. The difference is calculated for each bank in the sample for each quarter, and mean difference is aggregated by year. The mean difference per year is plotted as a percentage of outstanding loans. LINM is the allowance model estimated from the lasso model that takes the limited information as the input.

the weights are large by exploiting variations in banks' losses (section 1.6.2), and 3) assessing whether using aggregate data in the analysis plays a role in the performance difference (section 1.6.3). Finally, in section 1.6.4 I discuss why is it that the managerial ILM allowance estimates consistently underperform, and I explain the performance difference between the models by focusing on managerial incentives.

1.6.1 Examining the Weights on Input Variables by the LINM and ILM

The statistical tests to examine the performance difference between the LINM and the ILM models in this paper use the RMSE. To examine the differences in the RMSE arising from the statistical models, I explicitly consider the coefficients of the variables from the underlying LINM and ILM. The purpose of this analysis is to identify significant differences in the weights that each model assigns to the input variables, and the changes in these weights over time. I use the coefficients generated from the LINM model as its weights. For the coefficients of allowances under the ILM, I fit an OLS regression using the same set of variables from the limited information model on the recognized allowances as the dependent variable. I generate the coefficients estimated on various windows of my sample data split by year. Hence, this approach allows me to compare the coefficients directly to understand if the ILM systematically under-weights variables in the model and if managers ignore information they already possess.

The coefficients from the LINM lasso model are reported in Table A.4 and the coefficients from the ILM model are reported in Table A.5. I join the two tables and present them in Table 1.8. The columns in the tables refer to the window used to estimate the models. For example, the column 1996 – 2005 refers to the coefficients from the model estimated in the period between 1996 and 2005. Focusing on the sample ending in 2008 because this year is when the LINM and ILM models begin to diverge, I compare the magnitudes of the weights assigned to the variables.

Table 1.8 suggests that the outperformance of the LINM is driven by two effects. First, the LINM model consistently assigns larger weights to the 30 days past due, 90 days past due, and nonaccrual variables. In contrast, the ILM model systematically under-weights these variables, as observed across all of the windows. In the LINM model, the weight in 2008 for 30 days past due

Table 1.8: *Coefficients from the LINM and ILM Model of Sample Banks*

	1996 – 2002		1996 – 2005		1996 – 2006		1996 – 2007		1996 – 2008		1996 – 2010	
	LINM	ILM	LINM	ILM	LINM	ILM	LINM	ILM	LINM	ILM	LINM	ILM
30 days PD	0.077	0.028	0.090	0.030	0.089	0.029	0.095	0.028	0.129	0.030	0.167	0.032
90 days PD	0.183	0.013	0.268	0.013	0.304	0.017	0.324	0.020	0.355	0.024	0.388	0.026
Nonaccrual PD	0.223	0.088	0.274	0.102	0.291	0.104	0.318	0.104	0.406	0.108	0.522	0.126
30 days PD (t-3)	0	−0.013	−0.011	−0.017	−0.009	−0.017	−0.006	−0.016	−0.017	−0.016	−0.023	−0.016
90 days PD (t-3)	0	−0.007	−0.034	−0.007	−0.052	−0.010	−0.058	−0.012	−0.058	−0.015	−0.064	−0.016
Nonaccrual PD(t-3)	0.034	−0.030	0.051	−0.038	0.045	−0.041	0.039	−0.041	0.044	−0.039	0.002	−0.048
Loan to Asset	0	−0.015	0.004	−0.015	0.005	−0.015	0.007	−0.015	0.013	−0.014	0.010	−0.014
Securities to Asset	0	0.002	−0.000	0.002	−0.001	0.002	−0.001	0.002	−0.000	0.002	−0.005	0.001
Pct RE Loans	−0.007	0.000	−0.018	0.002	−0.018	0.001	−0.017	−0.000	−0.014	−0.001	−0.009	−0.002
Pct CI Loans	0	−0.003	−0.006	−0.003	−0.008	−0.004	−0.008	−0.004	−0.009	−0.005	−0.010	−0.006
Log Assets	0.088	−0.156	0.029	−0.149	0.077	−0.150	0.264	−0.147	0.600	−0.152	0.832	−0.138
Net-Chargeoffs$_t$	0	−0.033	0.133	−0.049	0.157	−0.046	0.160	−0.044	0.211	−0.026	0.334	0.003
Net-Chargeoffs$_{t-3}$	0.007	−0.005	0.119	0.000	0.129	0.007	0.127	0.013	0.180	0.019	0.253	0.057
Net-Chargeoffs$_{t-6}$	0.006	0.053	0.105	0.056	0.112	0.058	0.110	0.063	0.108	0.071	0.152	0.103
Net-Chargeoffs$_{t-9}$	0	0.074	0.053	0.075	0.062	0.075	0.057	0.080	0.044	0.081	0.073	0.109

This table presents the coefficients of the LINM model, and from the ILM model from implicit regression on allowances recognized by banks. The table presents combined coefficients from Table A.4 Table A.5. The columns represent coefficients from LINM and ILM models run between the periods 1996 – 2002, 1996 – 2005, 1996 – 2006, 1996 – 2007, 1996 – 2008 and 1996 – 2010 for sample banks. See Section A.2.1 for variable definitions.

is 0.129, the weight for 90 days past due is 0.355, and the weight for nonaccrual is 0.406. The corresponding weights in the ILM are 0.030, 0.024, and 0.108.

Second, the LINM continuously recalibrates and updates to incorporate the underlying information while predicting allowances, unlike the ILM. This aspect is reflected in the difference in the magnitudes of the 30 days past due, 90 days past due, and nonaccrual variables between the LINM and ILM. This difference gradually increases across the period. Furthermore, the magnitudes of the coefficients of LINM for the 30 days past due, 90 days past due, and nonaccrual variables are higher in 2008 relative to 2002. To provide economic intuition, the sum of the absolute differences between the coefficients of the three variables in 2002 is 0.35 percent of loans. The sum of the absolute difference doubles in 2008 to 0.73 percent of loans.[30] Using these variables alone, LINM would predict that banks would have increased their allowances by 0.73 times their loans in 2008.

The larger coefficients in 2008 for LINM occur because the model uses the experience in the period 1996 – 2008 to assign weights to the relevant input variables. An alternative explanation is that the LINM assigns larger weights to more recent data, while the ILM data assigns larger weights to historical data and smaller weights to recent data, which could improve its relative performance. To test this alternative explanation, I estimate the LINM and ILM models using only the period 1996 – 2002 and predict the allowance in the period 2005 – 2012. I then compare the performance of the models relative to the full rolling window case. The cumulative RMSE for the models are presented in Figure A.1. Here I report the RMSE beginning only in 2005. LINM_2002 refers to the lasso model estimated using the sample in the period 1996 – 2002, while ILM_2002 refers to the ILM model estimated in the period 1996 – 2002. I find that the LINM consistently outperforms the LINM_2002. The difference between the LINM_2002 and the LINM is attributed to the LINM model's incorporation of current information and experience to update the coefficient estimates of the input variables, which improves its performance. However, the performance of the ILM_2002 does not improve, and continues to perform poorly relative to the ILM, and in turn the LINM. The ILM model seems to partly ignore both recent and historical experience, as reflected

[30] This number for 2008 is calculated as the sum of the following differences – for the 30 days PD: $(0.129-0.030)$, 90 days PD: $(0.355-0.024)$, and Nonaccrual PD: $(0.406-0.108)$.

in the under-weightings. Thus, the test results rule out the alternative explanation for the poor performance of the ILM.

1.6.2 Exploiting Variation in Banks Operating in States with Severe and Less Severe Crises

I now identify a setting where the weights assigned by the LINM and ILM models play a large role in driving a wedge between their relative performances. The findings from Figure 1.4 and Figure 1.5 reveal only modest differences in the LINM and the ILM models' performance until 2007. Their performance diverges beginning in the year 2008. To examine this divergence, I partition the sample of all county banks into a subsample of banks that operate in states with severe crises, and another subsample comprising banks with less severe crises.[31] The partition allows me to exploit variation and observe differences in banks depending on the severity of the loan defaults in the state. The idea is that the banks operating in the severe default states suffered larger losses or recognized lower allowances, or both, in the period of the study. Therefore, the weights assigned to the input variables by the LINM and ILM should have larger effects on the performance of the models. To study the performance, I fit the LINM and ILM models for the full sample of banks as in the main analysis.[32] I then decompose the RMSE from the LINM and ILM models (presented in the top panel of Figure 1.1) by using the estimates from the models to predict allowance, and estimate the RMSE separately for the subsamples of banks in states with severe and less severe crises.

Figure 1.14 presents the out-of-sample cumulative RMSE for the two subsamples and the full sample. The models used to calculate the RMSE are the LINM and ILM using the full sample. The

[31] Given the focus here is to understand the differences in 2008, I classify states that suffered severe banking failures during the mortgage crisis by considering states that had at least 10 bank failures in the period 2008 – 2011. The states that are in the category are AZ, CA, FL, GA, IL, MI, MN, MO, NV, and WA. This classification is same as the one used in Section 1.5.6. The rest of the states comprise the category of less crisis states. I use this classification for the entire period of my study.

[32] This is one of the differences between the analysis in this section compared to the analysis in Section 1.5.6. In this section, I estimate using all sample banks, while in Section 1.5.6, I estimate separate model for the subsample. The other difference between this section and Section 1.5.6 is that the key focus of analysis in Section 1.5.6 is to compare the performances of the ILM, LINM, and the FINM model. But the focus in this section is to specifically understand the performance difference between the ILM and LINM.

RMSE for the sample of banks operating in the less severe states (LINM_Less and ILM_Less) and severe states (LINM_Sev and ILM_Sev) are also shown. There is a significant increase in the RMSE of ILM_Sev relative to the baseline ILM. The RMSE of ILM is 6.07, while the RMSE of ILM_Sev is 10.03. In addition, there is a decrease in the RMSE of ILM_Less relative to the ILM. The RMSE of ILM_Less is 2.17. Therefore, the RMSE of the ILM model is dominated by the RMSE of ILM_Sev. Hence, the findings suggest that the underperformance of the ILM relative to the LINM is driven by the performance of the banks operating in the severe crisis states.

Figure 1.15 presents the cumulative RMSE by year for the subsamples of banks in states with severe and less severe crises. This analysis is similar to the analysis underlying Figure 1.4 for the full sample case. The top panel is for the sample of banks operating in less severe crisis states, and the bottom panel is for more severe crisis states. The LINM_Less outperforms ILM_Less, and the LINM_Sev outperforms ILM_Sev. There is a striking contrast between the figures in the top panel and the bottom panel. The difference in the models for the less severe banks is smooth and the performance in 2008 is not pronounced. The bottom panel for severe banks reveals the spike beginning in 2008 for the ILM, similar to the main analysis with the full sample case. The underperformance of the banks in severe crisis states has a punitive effect on the performance of the ILM. Overall, the evidence in figures 1.14 and 1.15 suggests that the differences between the LINM and ILM, particularly beginning in 2008, are driven by banks operating in states with severe crises. The difference is attributable to the possibility that these banks possibly underprovisioned relative to the larger losses that they actually realized, as compared to the banks in less severe crisis states.

1.6.3 The Role of Aggregate Data in the LINM and ILM Models

The LINM model in this paper relies on using aggregate data on all county-level banks across the US to estimate allowances. One potential problem with this approach is that it may not be feasible in practice for a particular bank to use aggregate data across all banks. This is important because of the possibility that using the aggregate data from the sample banks drives the performance gains in the LINM. There are two reasons this concern is mitigated to a certain extent in my setting.

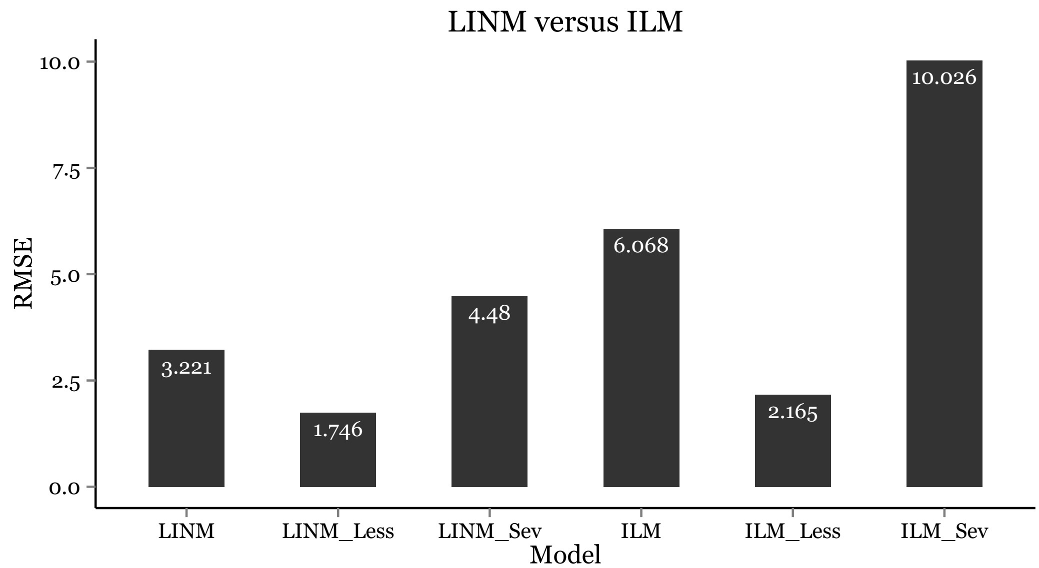
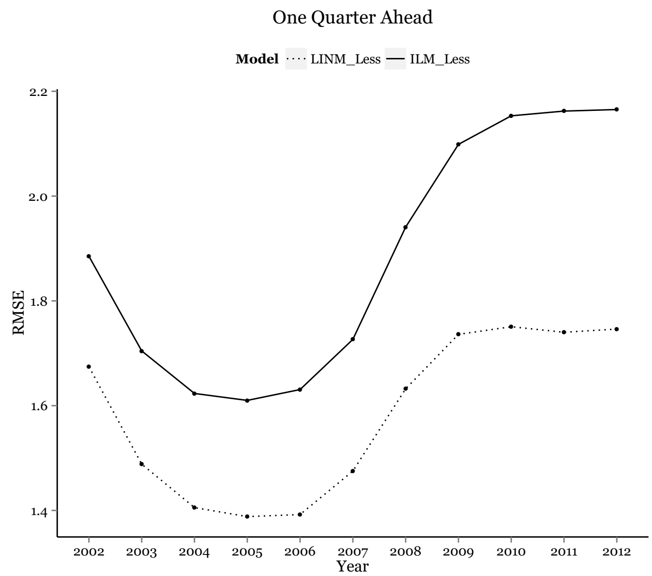


Figure 1.14: *LINM vs ILM Decomposed by Banks with Severe Banking Crisis States and Less Severe Banking Crisis States*

This figure presents evidence on the performance drivers of the LINM model relative to the current GAAP's ILM. The bars represent the cumulative root mean squared error for the models estimated and tested out of sample in the period 2002–2012. The overall sample of banks are split into banks that operate in US states that suffered severe banking crisis (Sev), and in states that suffered less severe banking crisis (Less). The severe states are identified by the ones that have more than 10 bank failures in the period 2008 – 2011. LINM is the root mean squared error of the predicted allowance from the lasso model that takes limited information as input, while ILM is the root mean squared error (RMSE) of allowance under current GAAP from the financial statements. The LINM_Sev is the root mean squared from the LINM model for banks in the severe crisis states, and LINM_Less is the root mean squared from the LINM model for banks in less severe crisis states. The ILM_Sev is the root mean squared from the ILM model of banks in the severe crisis states, and ILM_Less is the root mean squared from the ILM model of banks in less severe crisis states. Refer Figure 1.4 for further discussion on the models.



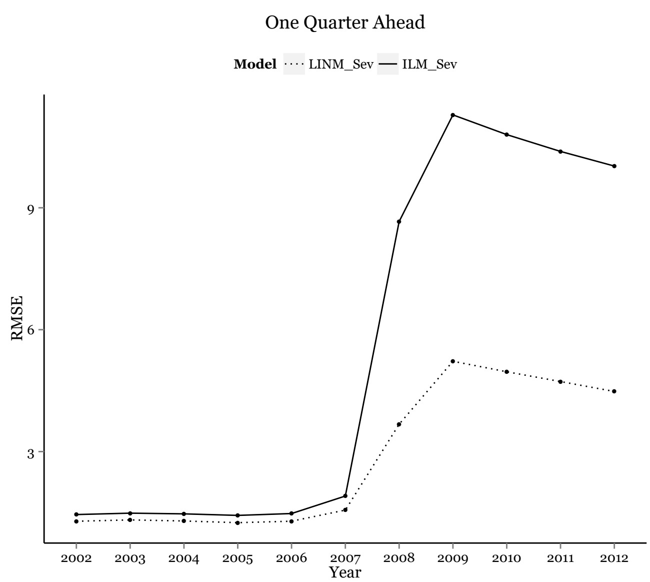


Figure 1.15: *Allowance Model Cumulative Performance By Year in US States with Less Severe and Severe Banking Crisis*

This figure presents evidence on the performance drivers of the LINM model relative to the current GAAP's ILM in their accuracy in predicting future losses. The overall sample of banks are split into banks that operate in US states that suffered severe banking crisis (Sev), and in states that suffered less severe banking crisis (Less). The severe states are identified by the ones that have more than 10 bank failures in the period 2008 – 2011. LINM is the root mean squared error of the predicted allowance from the lasso model that takes limited information as input, while ILM is the root mean squared error of allowance under current GAAP from the financial statements. The LINM_Sev is the root mean squared error from the LINM model for banks in the severe crisis states, and LINM_Less is the root mean squared from the LINM model for banks in less severe crisis states. The ILM_Sev is the root mean squared from the ILM model of banks in the severe crisis states, and ILM_Less is the root mean squared from the ILM model of banks in less severe crisis states.

First, the data from bank balance sheets used in the study are public and accessible to all banks. Second, it is reasonable to posit that through their auditors or regulators, bank managers receive information on the loan performance and allowance recognition of other banks.

Nevertheless, to address the concern directly, I repeat the analysis of the LINM and ILM by restricting them to all sample (one-county) banks within a specific geographic area. The focus is on all sample banks operating in a single state or in a cluster of states. The LINM and ILM models are then estimated by restricting them to only these banks, and the performances are compared in each specific case. If the results of this analysis, focusing on subsamples of banks, are consistent with the full sample case (with the LINM outperforming the ILM), then the analysis provides direct evidence that this paper's results are not driven by the use of aggregate data across all banks.

To test this premise, I focus on states that have a large number of sample banks. The states that fall in this category are Illinois and Minnesota. I also focus on states with the mean and median numbers of sample banks: Arkansas and New Jersey. Finally, I consider Maine, which has a small number of sample banks. In addition, to include a larger subsample of banks, the analysis is repeated by focusing on clusters of states – regions – designated by the Bureau of Economic Analysis (BEA). The regions are New England, Mideast, Great Lakes, Plains, Southeast, Southwest, Rocky Mountain, and Far West.[33] Finally, I also include banks in California, a state that is classified as a severe crisis state. The selection of the specific states and the regions is for the sake of parsimony in reporting the findings.

Figure 1.16 and Figure 1.17 present the cumulative RMSE estimated on sample banks restricted to the corresponding single state. The cumulative model performances for sample banks restricted to BEA regions are presented in Figure A.2 and Figure A.3. For example, the RMSE for sample banks in Minnesota is LINM: 1.79 and ILM: 2.28. The RMSE of sample banks in the BEA-designated Far West region is LINM: 2.56 and ILM: 3.56. Across all the states and regions in the analysis, the

[33]The BEA regions and the states that capture the areas are as follows: New England Region – Connecticut, Maine, Massachusetts, New Hampshire, Rhode Island and Vermont; Mideast Region – Delaware, DC, Maryland, New Jersey, New York, Pennsylvania; Great Lakes Region – Illinois, Indiana, Michigan, Ohio, Wisconsin; Plains Region – Iowa, Kansas, Minnesota, Missouri, Nebraska, North Dakota, South Dakota; Southeast Region – Alabama, Arkansas, Florida, Georgia, Kentucky, Louisiana, Mississippi, North Carolina, South Carolina, Tennessee, West Virginia; Southwest Region – Arizona, New Mexico, Oklahoma, Texas; Rocky Mountain Region – Colorado, Idaho, Montana, Utah, Wyoming; Far West Region – Alaska, California, Hawaii, Nevada, Oregon, Washington.

LINM model continues to outperform the ILM. The RMSE of the LINM is lower than the RMSE of the ILM. There are differences in terms of the magnitude of the cumulative RMSE across the states. Overall, the evidence in the figures from across the model performances is consistent with the hypothesis that the LINM outperforms the ILM even when smaller subsamples are considered, and rules out the hypothesis that the results are driven by using a large sample of banks.

1.6.4 Agency Issues in Recognizing Allowances

In this section, I consider why is it that the LINM brings performance gains relative to the estimates of the managers. To investigate this inefficiency, I focus on the incentives for managers to systematically understate the allowances as observed in the ILM. Note that the analysis in the paper thus far does not explicitly model the issue of agency in banks' allowance recognition. Therefore, I introduce incentives to the LINM and ILM models and check for differences in the models' accuracy and the magnitude of the allowances predicted from the models. I focus on regulatory capital management incentives in banks. I pick capital management for two reasons: first, due to my sample's limitations in identifying variations in other managerial incentives, and second, to build on prior research that studies the relationship between allowance recognition and bank capital (see Ryan, 2012, chapter 3).

To investigate whether managerial incentives to recognize allowances are different for well capitalized relative to weakly capitalized banks, I exploit variation in the sample of large banks discussed in Section 1.5.7. The sample is split into a subsample comprising banks whose capital ratios are above the median capital ratio in the year-quarter (well capitalized), and another subsample that comprises banks that are below the median capital ratio in the year-quarter (weakly capitalized). I then estimate the LINM and ILM separately for each of the two subsamples.

The model performance for the below median capital subsample is presented in the top panel of Figure 1.18 and the model performance for the above median capital subsample is presented in the bottom panel of Figure 1.18. The figures show the cumulative RMSE from each subsample in the period 2002 – 2012. In each of the subsamples, the LINM continues to outperform the ILM with a lower RMSE. The RMSE for the above median capital banks are LINM: 2.10 and ILM:

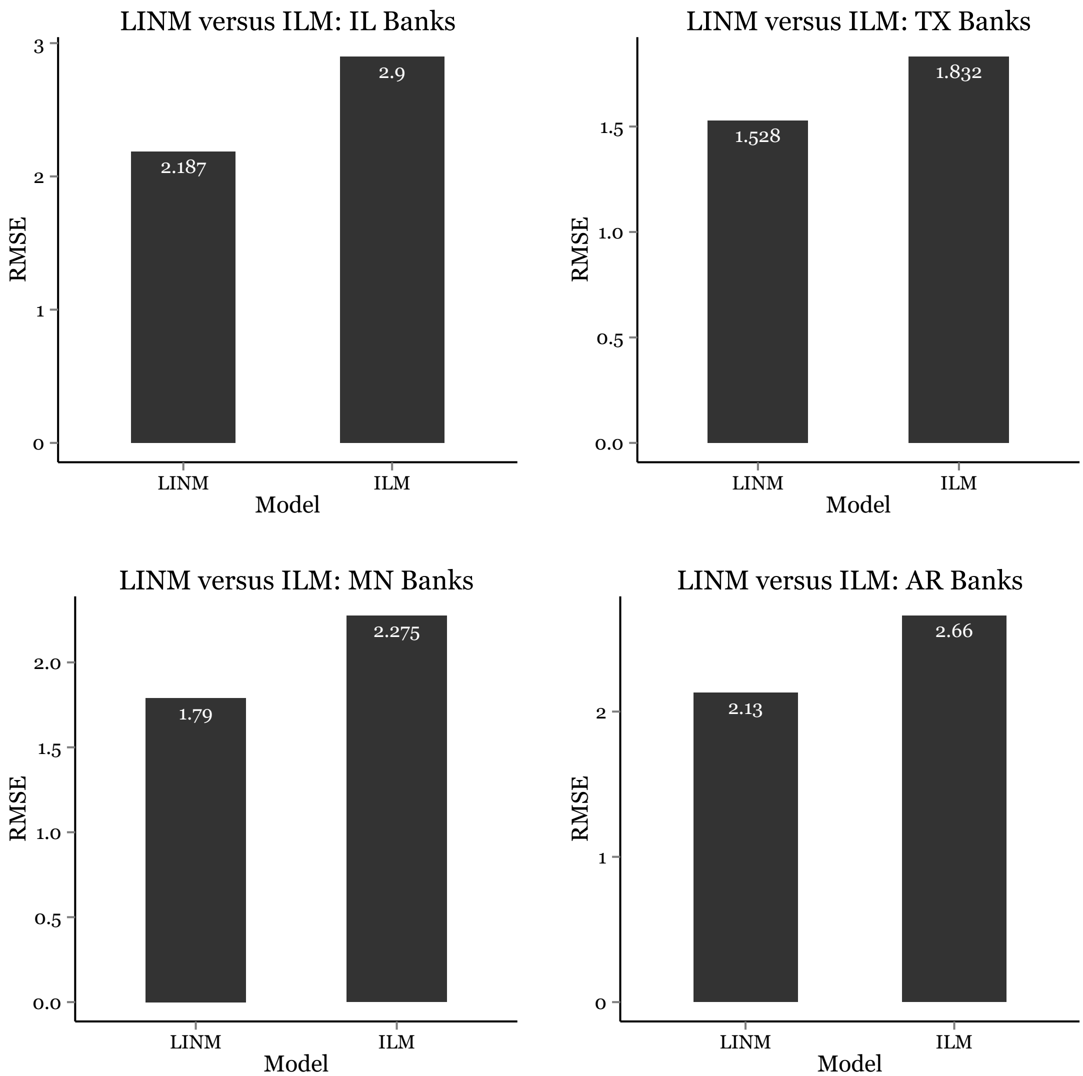


Figure 1.16: *LINM vs ILM for sample county banks operating in states*

This figure presents evidence on the difference in model performance based on sample banks that operate in particular states. The figures presents the performance of the LINM model predicted allowance relative to the current GAAP's ILM in their accuracy in predicting future losses. It shows the cumulative root mean squared error for the model estimated and tested out of sample in the period 2002–2012. LINM is the root mean squared error of the predicted allowance from the lasso model that takes the limited information as input. ILM is the root mean squared error of allowance under current GAAP from the financial statements. Each bargraph represents results from separate model estimations restricting on banks operating in the corresponding states.

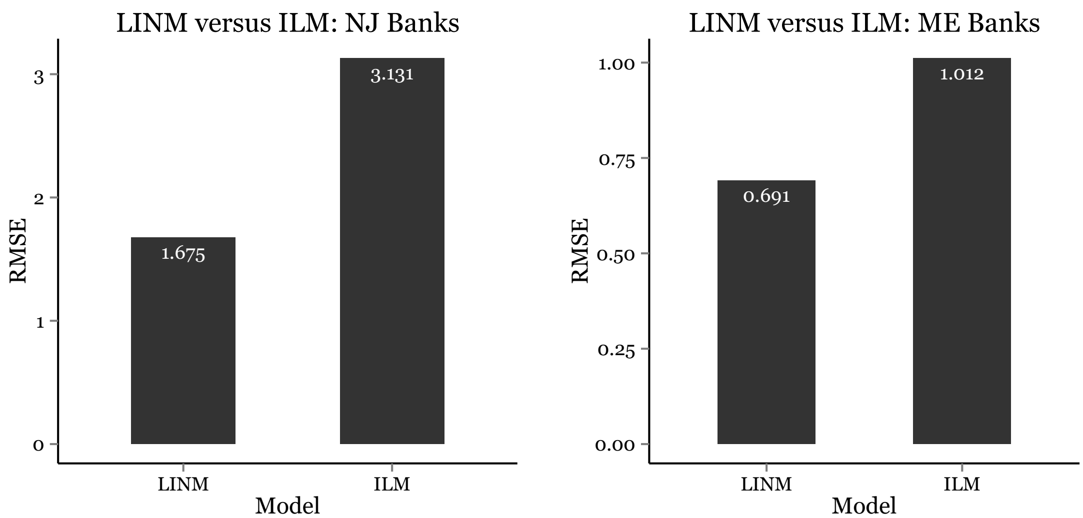


Figure 1.17: *LINM vs ILM for sample county banks operating in states*

This figure presents evidence on the difference in model performance based on sample banks that operate in particular states. The figures presents the performance of the LINM model predicted allowance relative to the current GAAP's ILM in their accuracy in predicting future losses. It shows the cumulative root mean squared error for the model estimated and tested out of sample in the period 2002–2012. LINM is the root mean squared error of the predicted allowance from the lasso model that takes the limited information as input. ILM is the root mean squared error of allowance under current GAAP from the financial statements. Each bargraph represents results from separate model estimations restricting on banks operating in the corresponding states.

2.92. The RMSE for the weakly capitalized (below median) banks are LINM: 1.88 and ILM: 2.89. There are no significant differences in the RMSE and model accuracy across the two subsamples. However, the difference in RMSE between LINM and ILM in the subsample of banks that are below the median capital is slightly higher compared to the difference in RMSE between LINM and ILM in the subsample of banks that are above the median.

The magnitude of the difference in the allowances, as a percentage of loans, from the models for the two subsamples is presented in Figure 1.19. The top panel is the difference for banks whose capital is below the median, the weakly capitalized banks, while the bottom panel is the difference for banks whose capital is above the median. The mean bank in the well capitalized bank subsample would have increased its allowance by 14%, the weakly capitalized subsample would have increased its allowance by 28%. This significantly larger increase for weakly capitalized banks is suggestive of their underprovisioning. In addition, the peak in the difference in magnitude for the below median capital firm is at 1.8 percent of loans, while the peak for the above median capital firm is about 1.2 percent of loans. The mean magnitude difference for the below median capital banks in the period 2007 – 2012 is 0.62 percent and the mean for the above median capital banks in the period 2007 – 2012 is 0.94 percent.

Overall, these findings – of a larger increase in allowances as predicted by the LINM for the weakly capitalized banks, along with a slightly higher difference in accuracy between the LINM and ILM models for these banks – provide evidence for the difference in provisioning between the subsamples. The evidence suggests that weakly capitalized banks exercise discretion and underprovision relative to well capitalized banks.

1.7 Conclusion and Suggestions for Future Research

This paper contributes to our understanding of the role of information in financial reporting and transparency. I examine this notion in the context of how accounting rules influence bank loan loss allowances, where a longstanding debate exists among standard setters and regulators about the best way to account for loan losses. Current GAAP's incurred loss model delays loss recognition until a probable threshold is met. This aspect has been criticized for limiting bank's

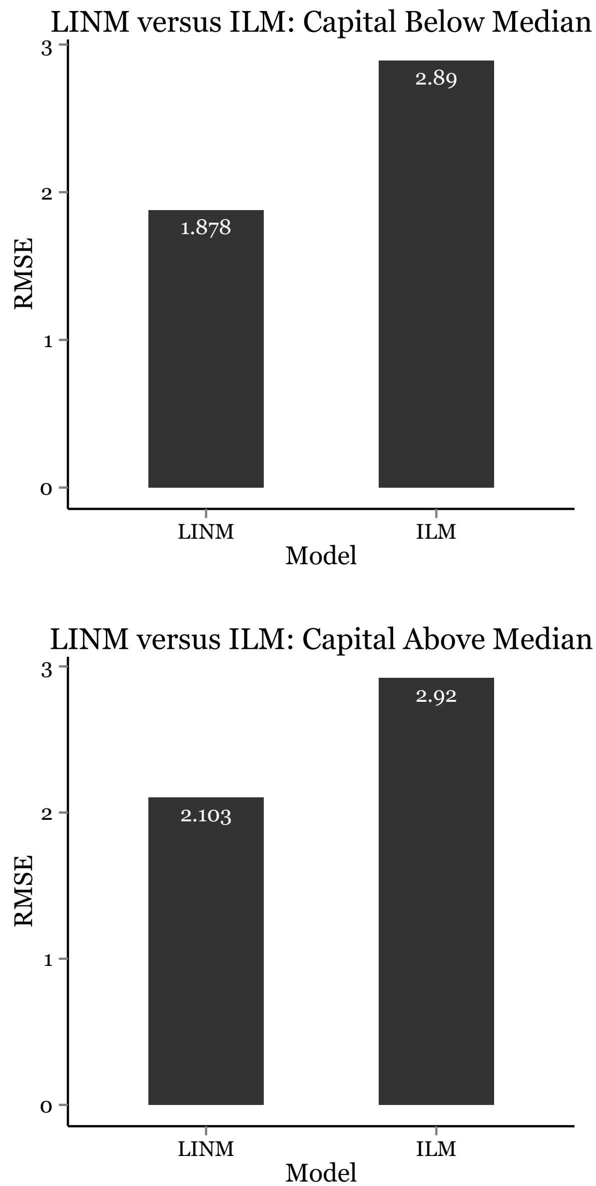
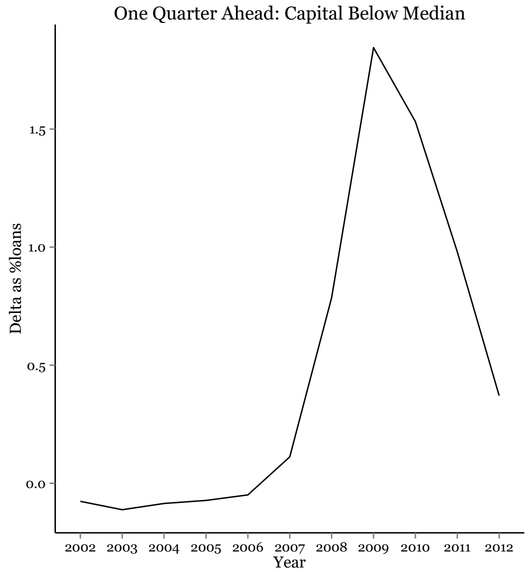


Figure 1.18: *Summary of Allowance Model Performance for Large Banks that are Above and Below Capital*

This figure presents evidence on the difference in model performance based on the degree of capital constrains for large banks. The figures presents the performance of the model predicted allowance relative to the current GAAP's ILM in their accuracy in predicting future losses. It shows the cumulative root mean squared error for the model estimated and tested out of sample in the period 2002–2012. LINM is the root mean squared error of the predicted allowance from the lasso model that takes the limited information as input. ILM is the root mean squared error of allowance under current GAAP from the financial statements. The sample of large banks are split based on if the bank's capital ratios are above or below median. The models are estimated separately for the sample above the median and below the median.



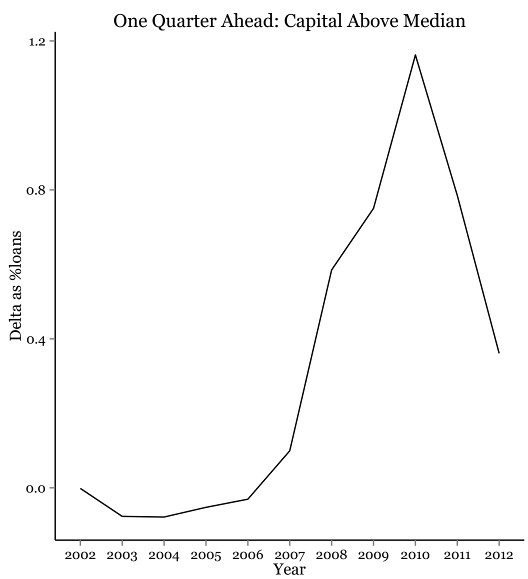


Figure 1.19: *Difference between Predicted Allowance and Actual allowance for Large Banks Based on their Degree of Capital Constrain*

This figure presents magnitude difference between the allowance estimated from the LINM and actual loan loss allowance from the bank's financial statements. The analysis is based on the degree of capital constrains for large banks. The sample are split based on if the bank's capital ratios are above or below median. The models are estimated separately for the sample above the median and below the median. The difference, as a percentage of loans, is calculated for each bank in the sample for each quarter, and mean difference is aggregated by year. LINM is the allowance model estimated from the lasso model that takes the limited information as the input.

ability to recognize losses that are expected, but have not yet met the threshold. The mortgage crisis underscored some of these concerns, and the FASB responded by issuing a new accounting standard CECL, that would alter the information set banks can use in estimating allowances to include forward-looking information.

I examine whether it is possible to construct a predictor of future loan losses that outperforms the incurred loss model by using the same information that is typically allowed under the rule. I then examine the impact of expanding the information set to include data of the kind proposed in the CECL rule (The CECL rule will be effective beginning December 2019). I find that it is indeed possible to construct a prediction model – the limited information model – that performs better than the incurred loss model, without having to expand the information set beyond what is already used under the incurred loss model. I also find that expanding information to this model provides no significant performance gains. I document that using the predicted allowances from the limited information model, banks would have recognized higher allowances before entering the financial crisis. I also find that the performance gains from the limited information model is consistent for a sample of large banks.

I then examine the factors that drive the outperformance of the limited information model relative to the ILM. I find that the limited information model assigns larger weights to input variables, and continuously recalibrates incorporating the underlying information, unlike the ILM. I find that these translate to large differences in performance, particularly for banks that operated in states with severe crisis. In examining the incentives of managers to understate the losses in the ILM, I find evidence consistent with that weakly-capitalized banks underprovision, relative to the well-capitalized banks.

While a large stream of accounting and banking research examines the factors that influence bank loan loss accounting, there is a paucity of empirical research that examines the rules and its implementation. I find that even with a simpler and eminently feasible approach, implemented from the machine learning literature in lasso outperforms current GAAP.

I further use the CECL rule change as a setting to understand on two aspects of the rule, the discretion allowed, and the use of the information. My findings suggest that using broader information in allowance estimates is beneficial, but that primary benefits are from how the discretion is exercised in using the information, rather than the availability of broader information. I argue that this finding raises the question as to whether discretion improves accounting quality in my setting, and contributes towards the debate in standard setting, and financial reporting over affording managers discretion versus a uniformity in the rules.

The findings in this study suggests several avenues for future research. One natural question is to understand the form of inefficiency in the accounting rule. It would be an interesting study to extend the analysis to understand if the outperformance of the limited information model is driven by the biases from the models used by the banks, or other incentives of bank managers in applying the rule. Another area is to explore whether bank managers react to the limitations in the accounting rules through real actions such as raising capital or reducing risk or through other means. Finally, it would be interesting study to examine if the market understands the frictions arising from current GAAP, and if there are possible mispricing of risks. Understanding these questions would further increase our grasp of bank accounting choices, financial reporting, and more broadly to standard setting.

Chapter 2

What Else do Shareholders Want? Shareholder Proposals Contested by Firm Management [1]

2.1 Introduction

Shareholder proposals provide a means for investors to communicate their demands to a firm's management and board of directors. Such proposals, allowed under the Securities Exchange Act of 1934, offer shareholders the opportunity to propose changes that other investors can vote upon. A growing body of research indicates that shareholder proposals provide an effective tool to promote changes in compensation policy, firm strategy, and governance (Yermack, 2010; Ferri, 2012). Although shareholders are given the opportunity to present proposals to enact changes, this is not an unencumbered right. Management, with permission of the Securities and Exchange Commission (SEC), can exclude shareholder proposals from appearing on the proxy statement.

We examine whether management's desire to exclude shareholder proposals from the proxy potentially encumbers shareholders' rights to influence firm policy. On one hand, managers may seek to only exclude those proposals that represent infeasible ideas or the personal interests of

[1] Co-authored with Eugene Soltes and Suraj Srinivasan

minority shareholders. Excluding these proposals can be viewed as increasing the efficiency of the proxy voting process by not presenting frivolous matters to shareholders. On the other hand, managers may seek to exclude proposals that pose threats to their own interests. If so, seeking their exclusion could potentially be an expression of managerial entrenchment and counter to the interests of shareholders.

We hand-collect all proposals that managers seek to exclude from the proxy from 2003–2015. We find that managers often seek to exclude shareholder proposals from the proxy. Nearly five thousand proposals, or nearly 40%, of all proposals received during our sample period are contested by management. These proposals cover a wide range of issues including executive compensation, antitakeover measures, voting procedures, environmental issues, and social policy. The SEC allows firms to exclude many of the proposals that managers contest. Specifically, 73% of all proposals that managers seek to exclude from the proxy are allowed by the SEC (i.e., SEC provides the firm a "no action" opinion letter). This suggests that the SEC's criteria of allowed exclusions, which varies and changes over time, significantly affects what shareholders are given the opportunity to vote on.

Some firms are more inclined than others to seek exclusion of shareholder proposals. Firms that are larger, have worse performance, and have less institutional shareholders are more likely to contest the proposals they receive. There is also the tendency for management to behave similarly over time. If a firm contests a proposal in the prior year, they are more inclined to contest a proposal in the subsequent year.

For the 27%, or $1,332$ shareholder proposals, that managers cannot exclude from the proxy, managers have a choice. They can either proceed with placing it on the proxy or they can seek to engage with the submitting shareholder to reach some compromise that would lead to the shareholder withdrawing the proposal before the vote. Our evidence supports the idea that managers often seek to exclude proposals that are not necessarily frivolous and are supported by a significant proportion of shareholders.

Regulation only requires that shareholders hold $\$2,000$ of stock or 1 percent of the share capital for at least one year to be eligible to create a proposal.[2] However, we find that the vast majority of shareholders who create proposals have considerably larger holdings. The median shareholdings of submitters whose proposals are contested have $39,000 in shares and the mean submitter has $10.7 million in share ownership (1.4% of outstanding shares). This skew reflects the holdings of larger institutional investors (e.g. pension funds, hedge funds, etc.) whose proposals are often contested. Thus, the shareholders whose proposals are contested are typically not marginal holders of the firm's securities.

More importantly, for the proposals that are contested, but the SEC does not offer an exclusion, we find that 17% of all contested proposals that are brought to a shareholder vote are approved by shareholders. By comparison, 25% of non-contested proposals that are placed on the proxy receive majority shareholder support.[3]Thus, proposals contested by management that eventually make their way to the proxy often gain broader shareholder support at a level comparable to non-contested proposals. Notably, even those contested proposals that fail to be approved still gain considerable support. Contested proposals that fail, gain on average an additional 21% percent of all shares outstanding in incremental support over the shares held by the submitter. This suggests that even those proposals that fail are not entirely frivolous given this magnitude of shareholder support. Together, this evidence suggests that managers often contest proposals that are supported by their shareholder base.

After contesting a proposal and even after receiving a no-action letter, managers still have an alternative to placing the proposal on the proxy. In particular, managers can negotiate with shareholders prior to a vote. These negotiations can lead to the withdrawal of the proposal by the submitting shareholder after the submitter is satisfied with the firms' actions. 23% of all contested proposals are eventually withdrawn by the submitting shareholder or simply implemented by the firm (i.e. in effect withdrawing the proposal). We find that managers are more willing to negotiate with shareholders with larger holdings and with institutional entities like pension funds

[2]Division of Corporation Finance: Staff Legal Bulletin No. 14.

[3]As another benchmark, between 1973 and 2004, less than 10% of proposals received majority support (Gillan and Starks, 2007).

and hedge funds. Mangers are also considerably more likely to conclude a private resolution with the shareholder once the SEC disallows the firm from excluding the proposal from the proxy.

Overall, our evidence is consistent with managers often seeking to exclude proposals that represent the interests of their shareholders. In 21% of proposals that managers sought to exclude but were not permitted to do so by the SEC, shareholders approve the proposal through a vote or the submitter withdraws due to implementation. Given that shareholder interests regularly seem to differ from that of management, this analysis suggests that which proposals are excluded plays an important role in determining governance outcomes. The SEC's selection process plays an important role in either facilitating or encumbering shareholder interests.

Our paper contributes to the corporate governance literature in several ways. An extensive literature has examined shareholder activism through the proxy (for reviews, see Karpoff, 2001; Yermack, 2010; Ferri, 2012). While the earlier literature reviewed in Karpoff (2001) suggested that shareholder proposals have limited impact, their effectiveness has become more significant in recent times (Ferri, 2012). Extant research on shareholder proposals ignores a significant portion of submitted proposals since they are excluded by companies from inclusion in the proxy. To our knowledge, our analysis is the first to examine contested proposals where management actively seeks to exclude shareholder proposals from the proxy. By understanding which proposals management seeks to exclude, we are able to better understand managers' desire to be receptive to shareholder interests.

Our analysis also contributes to understanding the changing nature of shareholder interests and the implementation of regulation around the proxy voting process. In an interpretive release issued by the SEC in 2009, staff from the Division of Corporate Finance noted that "over the past decade, we have received numerous no-action requests from companies seeking to exclude proposals relating to environmental, financial, or health risks . . . based on our experience reviewing these requests, we are concerned that our application of the analytical framework . . . may have resulted in the unwarranted exclusion of proposals" (SEC, 2009). This suggests that while some proposals appeared at one point as "fringe" proposals that could be rejected as simply being an individual grievance, regulators are now beginning to pay attention to these concerns. By showing that many of the proposals that are contested, and later excluded by the SEC, are made by large engaged shareholders, we illuminate the effect of regulators' decisions to potentially exclude certain types of proposals.

The rest of the paper proceeds as follows. Section 2.2 discusses the shareholder proposal process and offers examples of the shareholder contestation process. Section 2.3 examines the firms and proposals that are contested and investigates the determinants of the process. Section 2.4 the implications of our analysis within the broader governance context. Section 2.5 concludes.

2.2 Institutional Background

2.2.1 Shareholder Proposal Process

Shareholders have a variety of mechanisms to mitigate agency conflicts created by the separation of ownership and control. Two options available to investors are to sell their shares thereby exiting the firm or initiating a takeover to gain complete control (Admati and Pfleiderer, 2009; Parrino *et al.*, 2003). Along the spectrum of alternative options lies shareholder engagement of various types though which investors can participate in the company's strategic direction. For instance, individual shareholders can press for corporate reforms by negotiating with the management privately (Carleton *et al.*, 1998; Strickland *et al.*, 1996; Becht *et al.*, 2007). Management may not be receptive to these ideas which can hinder the success of such a dialogue. Consequently, the shareholder may be unable to reach a satisfactory private resolution.

An alternative mechanism arises under Rule 14a–8 that was first promulgated by the SEC in 1942 (Palmiter, 1993; Prevost and Rao, 2000; Brown, 2016).[4] Shareholders can create proposals that are placed on a firm's proxy statement which is distributed to shareholders before its annual

[4]The impetus for the SEC's shareholder proposal rule arose from the efforts of John and Lewis Gilbert, two prominent activists who sought to motivate greater shareholder engagement in corporate governance matters. In 1939, the Gilberts' sought to amend the bylaws of Bethlehem Steel by seeking shareholder election of the firm's auditors. Mangers ignored the Gilbert's motion when they sent out the proxy which ultimately drew the attention of the SEC. As the SEC explained in 1940, "the Commission has been seriously concerned regarding the responsibility of corporate management to communicate to security holders information with respect to matters which minority groups have indicated will be brought up for action at a proposed meeting."(SEC Exchange Act Release 2376). The SEC later adopted its initial shareholder proposal rule in 1942 specifically giving shareholders access, under certain conditions, to the proxy. See Fisch (2008).

meeting. Under Rule 14a-8 of the Securities Exchange Act of 1934, a shareholder that has held $2,000 worth of shares or 1% of market value of equity continuously for at least a year is allowed to include a proposal in the company's proxy for a vote at the annual meeting.

Shareholder proposals placed on a firm's proxy under Rule 14a–8 offer shareholders a direct opportunity to influence a firm's corporate policies (Thomas and Cotter, 2007).[5] Allowing proposals by shareholders helps alleviate the agenda setting problem where there is no other alternative but to support the decision of managers (Pozen, 1994; Ryan, 1988). Since the SEC expanded shareholder rights in the 1940's, shareholders have sought to offer proposals covering a wide range of corporate governance (e.g., compensation, anti-takeover provisions, declassifying boards, and supermajority requirements) and social issues (e.g., environmental policy, employment equality, and ethical conduct).[6]

While shareholder proposals can advance policies that are aligned with the interests of all shareholders, some shareholders may seek to co-opt this process as a means to further their own specific economic or personal agendas. For example, the animal-rights organization People for the Ethical Treatment of Animals (PETA) commonly utilized shareholder proposals in the 1980's to advocate against animal testing practices by household and cosmetic manufacturers. Advocacy groups, like PETA, often gain ownership of shares of companies they want to target from benefactors. These shares are then sold after the resolution of the proposal or retained for the purpose of bringing later resolutions to public attention. As this example suggests, some shareholders may invest in the firm for objectives other than long-term profit maximization.

To address this potential problem, one solution is to allow managers to exclude proposals that, in their opinion, would not be in the best interests of all shareholders. However, allowing such

[5] A large empirical literature discusses different areas of the voting of directors and executive compensation. This work includes Shivdasani and Yermack (1999), Gompers *et al.* (2003), Bebchuk and Cohen (2005), Faleye (2007), Subramaniam and Wang (2009), and Ferri and Sandino (2009).

[6] The seminal case that clarified the scope of the SEC with regard to shareholder voting rights was SEC v. Transamerica Corp., 163 F.2d 511, 518 (3d Cir. 1947). In this case, management sought to exclude a proposal since management saw it as in conflict with Delaware state law. In the only case of its kind (Fisch, 2008), the SEC brought suit against Transamerica to compel it to include the proposal. Ultimately, the Court of Appeals appealed to federal law and concluded that Congress did intend to give shareholders voting rights consistent with the SEC's rule. In particular, the court stated: "The power conferred upon the Commission by Congress cannot be frustrated by a corporate by-law." (163 F.2d 511, 518).

managerial discretion permits management to not only exclude proposals that are disadvantageous for the firm, but also bona fide proposals that most shareholders would support but which hinder managements' own interests. Such exclusions can create barriers that restrict shareholder influence on the firm.

In order to define manager's ability to restrict shareholder's access to the proxy, the SEC laid out conditions that managers could exclude proposals from the proxy in Rule 14a-8. Beginning in 1947, managers seeking to exclude a proposal from their firm's proxy were required to submit it to the SEC for a "no-action" review to exclude a proposal (Palmiter, 1993).

Over time, the SEC refined its guidance of which actions it would issue "no action" letters by creating criteria of the types of proposals firms may exclude from its proxy (with the current criteria described in detail in the next section). These criteria include such matters as the form of the proposal, whether it duplicates other proposals/previously implemented proposals, and whether the proposal is a matter of ordinary business operation. Despite the creation of this criteria, attorneys at the SEC are given considerable discretion in interpreting whether a proposal may be excluded (Steel, 2016).

As an example of how this discretion influences the type of proposals ultimately placed onto proxies, in 1992 the New York City Employees' Retirement System (NYCERS) submitted a proposal to Cracker Barrel requesting that the firm bar discretionary hiring practices on the basis of sexual orientation. The SEC responded with a no-action letter permitting the proposal to be excluded on the basis of the proposal addressing a matter of ordinary business operations of employment policy. But in 1998, attorneys at the SEC reversed this decision and indicated that such employment proposals would be permitted if they raised other substantial policy considerations (Stanton, 1999; Fisch, 2008).[7].

More recently, attorneys at the SEC allowed firms to exclude proposals regarding CEO succession because such decisions were perceived as being part of ordinary business matters. However,

[7] The Commission's change in opinion in this matter seems to have been motivated by earlier institutional and Congressional pressure. In 1996, as part of the National Securities Markets Improvement Act, Congress required the SEC to reexamine its shareholder proposal process to see if it adequately provided appropriate access. Many of those that provided input were critical of the SEC's application and interpretation of the shareholder proposal rules under 14a–8 (see SEC, 1997a)

attorneys at the SEC changed their position on such matters in 2009. After this point, attorneys at the SEC began viewing CEO succession as a governance matter which should not be excluded from the proxy.[8] As these example indicate, attorneys at the SEC often change positions on which types of proposals are excludable with little notice to shareholders or issuers (Palmiter, 1993; Choi, 2006).

If the SEC deems the proposal as one that fits the exclusion criteria, it issues the firm a "no action letter" saying that SEC staff would not recommend the Commission take enforcement action against the firm if it excludes the proposal from its proxy. While the no-action letter is not legally binding and could be challenged in federal court or by review of the full SEC Commission, in practice the no-action letter settles the dispute (Bartkus, 2016; Lemke, 1987; Nagy, 1997). By looking at all the determinations by the SEC in 2015, for instance, Steel (2016) finds no additional actions brought in federal court and notes that the no-action letter normally is "the end of the matter" due to the high costs of litigating in court and relatively limited upside from the perspective of the filer.

With some rare exceptions, shareholder proposals do not bind a firm's board to actually implement the proposal irrespective of the extent of shareholder support they generate in the proxy vote (Levit and Malenko, 2011). Binding proposals can be excluded by the company as they impinge on the boards' prerogative to conduct the business of the company.[9]

2.2.2 Lifecycle of a Shareholder Proposal

Once a shareholder submits a proposal, managers have the choice of including it in their next proxy statement for a vote or seeking exclusion with the SEC. If the firm seeks to exclude the proposal, it must file the reason(s) for its exclusion with the Division of Corporate Finance at

[8]"We now recognize that CEO succession planning raises a significant policy issue regarding the governance of the corporation that transcends the day-to-day business matter of managing the workforce. As such, we have reviewed our position on CEO succession planning proposals and have determined to modify our treatment of such proposals. Going forward, we will take the view that a company generally may not rely on Rule 14a-8(i)(7) to exclude a proposal that focuses on CEO succession planning." (SEC, 2009)

[9]In 1976, the SEC created the "precatory proposal" which addressed this limitation by framing proposals as recommendations rather than mandates (Exchange Act Release No. 34–12999, Nov. 22, 1976).

the SEC. Firms request permission to exclude proposals by asking the SEC for its assessment of whether it would take any action if the firm excluded a particular proposal. If attorneys at the SEC believe that managers' reason to exclude to the proposal falls within the acceptable exclusions, the SEC will issue a "no action" letter stating it will not take any action if the firm excludes the proposal.

Managers can seek to have shareholder proposals excluded from the proxy for a variety of eligibility and/or procedural reasons. Rule 14a–8 of the Securities Exchange Act of 1934 offers management reasons under which they can request to exclude a shareholder proposal from the proxy. Acceptable exceptions include violating a state or federal law, being part of the company's ordinary business operations, containing a special interest, duplicating another previously submitted proposal, or being substantially implemented. Table 2.2 provides a complete list of reasons that firms can seek exclusion under Rule 14a–8 of the Securities Exchange Act of 1934.

If the SEC attorneys disagree with firm management that the proposal does not sufficiently meet the conditions for exception, the SEC issues a letter noting that the SEC is unable to concur with the firm's desire to exclude the proposal. Without concurrence by the SEC, managers have two additional options at their disposal prior to placing the proposal on the proxy for a shareholder vote. First, management can submit further supporting statements to convince the SEC to reconsider its original decision. In this case, the SEC can either maintain or change its original decision. Alternatively, management can negotiate with the shareholder. If the shareholder is satisfied with their negotiations with management, the shareholder can withdraw the proposal. In doing so, the proposal is omitted from the proxy and hence not voted.

Appendix B.1 provides a timeline describing the process of submitting, contesting, and voting on shareholder proposals. Figure 2.1 provides an illustration of the complete shareholder proposal process.

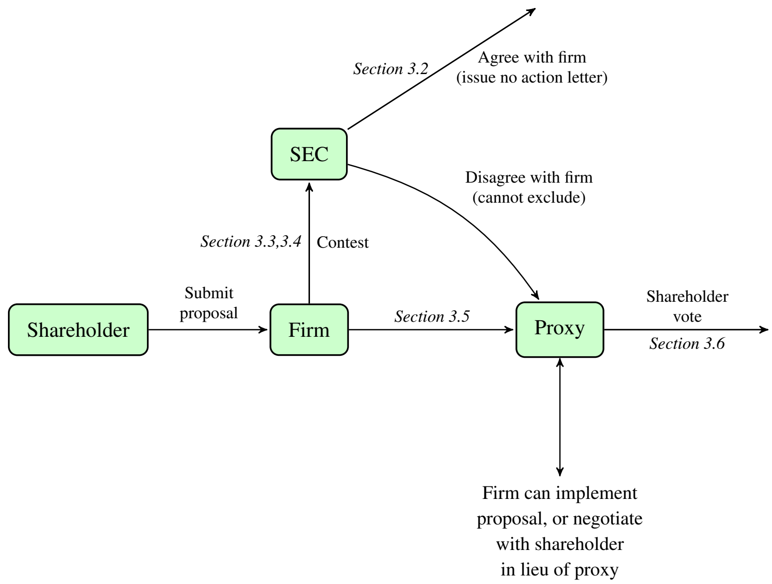


Figure 2.1: *Shareholder Proposal Process*

2.2.3 Examples of Contested Shareholder Proposals

We provide two examples to illustrate the process of resolving contested shareholder proposals. In the first example, the SEC allows the firm to exclude the proposal from the proxy. In the second example, the SEC does not permit the firm to exclude the proposal, but management implements the proposal prior to the proxy vote.

JP Morgan Chase– Risk Management Shareholder Proposal

In the final week of November 2010 the Sisters of Charity of Saint Elizabeth (hereafter *Sisters*), a faith based socially responsible investing group, submitted a proposal to JP Morgan Chase. They sought JP Morgan Chase to report to its shareholders "the risk management structure, staffing, and reporting lines of the institution and how it is integrated into their business model and across all the operations of the company's business lines." In addition, the Sisters also provided proof that the group held the required number of shares in JP Morgan Chase and their intention to hold these through until at least the annual meeting. The shareholder proposal letter described their interest in submitting the proposal as part of an attempt to help "restore confidence in the financial system."

JP Morgan Chase retained the law firm of O' Melveny and Myers LLP which sent a no–action request on the company's behalf to the SEC on January 10, 2011. JP Morgan Chase requested that it omit the shareholder proposal from its proxy materials by Rule 14a–8(i)(7) which allowed shareholder proposals that deal with a company's ordinary business operations and 14a–8(i)(10) which allowed exclusion if a firm had substantially implemented the proposal. The SEC responded on February 11, 2011, stating that it concurs with JPMorgan's view that it may exclude the proposal under Rule 14a-8(i)(7). Following this, JP Morgan Chase omitted the proposal from its proxy materials for its shareholder meeting in May 2011.

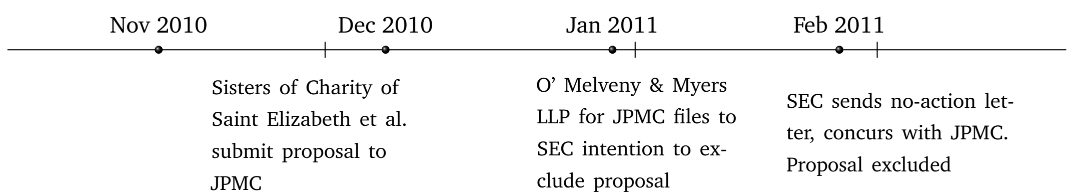


Figure 2.2: *Time Line for JP Morgan Chase – Risk Management Proposal*

3M – Animal Testing Shareholder Proposal

On November 21, 2005, two shareholders of 3M filed a proposal that sought for 3M to create an animal welfare policy that would reduce and replace the use of animals in testing and also provide better care for animals when used by the company and its contractors. The shareholders also designated the animal right group, People for the Ethical Treatment of Animals (PETA), as their legal representative. On January 6, 2006, 3M filed with the SEC their intention to omit the shareholder proposal from its proxy. 3M argued in its letter that the proposal was already substantially implemented, that the proposal was vague and indefinite, and that it was beyond 3M's power to implement more fully. It requested that the SEC staff concur with them and not recommend enforcement action if 3M excluded the proposal from its 2006 proxy.

On January 27, 2006, PETA filed a response with the SEC to 3M's arguments to omit the proposal rejecting all three of 3M's reasons for exclusion of the proposal. On March 10, 2006, the SEC responded, agreeing with PETA, and rejecting 3M's reasons to omit the proposal. Subsequently, in March 2006, 3M sent out its proxy statement for its annual meeting and included PETA's shareholder proposal regarding the animal welfare policy. The board recommended shareholders vote against the proposal. Management of 3M argued that the proposal was unnecessary since, in their opinion, they had already sufficiently implemented the animal welfare policy supported by PETA.

On May 8, 2006, the day before 3M's annual meeting, PETA issued a press release that stated that PETA had successfully negotiated with 3M and that 3M had implemented the desired animal welfare policy. 3M posted its animal welfare policy online and included measures to issue an annual report in compliance with animal welfare policy. The director of PETA's regulatory testing said: "we are delighted with 3M's response to our proposal." Despite the fact that it was never voted on, from PETA's standpoint, its proposal was successfully adopted and implemented by 3M.

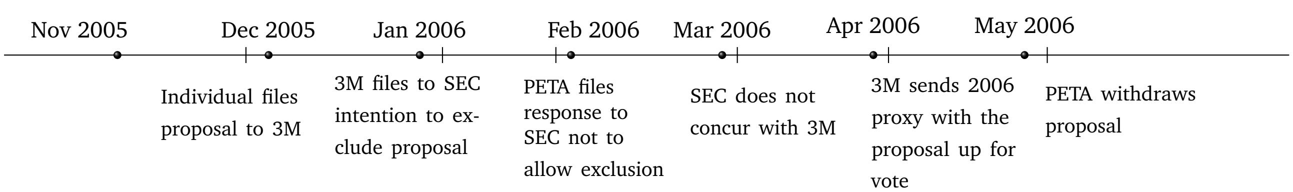


Figure 2.3: *Time Line for 3M – Animal Testing Shareholder Proposal*

2.3 Examination of Contested Shareholder Proposals

In Figure 2.1, we graphically describe the process of contesting and voting on shareholder proposals. Within the figure, we show the relevant section within Section 3 that examines the specific portion of the decision process.

2.3.1 Contested Shareholder Proposal Sample

We examine contested proposals issued under Exchange Act Rule 14a from 2003–2015. To find contested shareholder proposals prior to October 2007, we utilize the legal database IntelliConnect by Wolters Kluwer to manually collect the proposals and the related no-action letters. After October 1, 2007, we collect these letters online at the Division of Corporation Finance at the SEC.

For each contested proposal, we acquire the original shareholder proposal, correspondence between the shareholder, firm, and SEC, and the final opinion letter drafted by the SEC attorney. We manually code each proposal along the following dimensions: the firm name to which the proposal was submitted, the proponent, the type of proposal, the proposed reason(s) for exclusion by the firm, and the final decision of the SEC. We classify proposals by topic matter along the lines of prior research (Gillan and Starks, 2000; Gordon and Pound, 1993; Renneboog and Szilagyi, 2011; Thomas and Cotter, 2007). For the proposals that the SEC does not allow exclusion of, we track the proposal to determine whether it was voted on or withdrawn. We utilize the ISS Voting Analytics database to find proposals voted on at shareholder meetings and the results of those votes. To find proposals that were withdrawn, we read company proxies, search the Factiva news database, and review firm websites to ascertain each proposal's outcome.

For each proposal that is contested, we manually collect the number of shares held by the proponent from documents submitted to the firm by the shareholder along the proposal. For 77% $(3,771)$ of the shareholder proposals contested by management we are able to acquire the amount of ownership. Missing observations arise for three reasons. First, some proposals (2.1% of the sample) say that the letter is attached, but the letter is not available (presumably due to errors in the scan of the document by the SEC). Second, some proposals (10.2% of the sample) simply assert that the shareholder has the minimum $\$2,000$ in shareholdings, but it does not quantify the exact number of shares. Finally, for some other proposals (10.4% of the sample) the information is simply omitted from the original proposal sent by the shareholder.[10]

To supplement the contested proposals data, we utilize CRSP and Compustat for accounting and market pricing data. We find institutional ownership from the Thomson Reuters institutional database and board data from BoardEx. The table descriptions provide details on the particular definitions for each variable utilized in the analysis.

2.3.2 Frequency of Contested Proposals

We present in Table 2.1 statistics on the set of 14a–8 shareholder proposals that shareholders sought to include in the proxy in our same period 2003–2015 (i.e., including those not contested by the firm). Of the 12,475 proposals submitted in this time, 4,852 (or 39 percent) were contested by firms, suggesting that contesting shareholder proposals is not an unusual phenomenon. Firms are generally successful in this effort, with the SEC allowing exclusion of over 72 percent (3,520) of the proposals. But in a significant number (1,332) of cases, firms are forced to include the shareholder proposals that they initially sought to exclude. Our analysis focuses on these proposals. The proportion of shareholder proposals contested by firms has remained relatively stable over the sample period from 2003-2015. In this time, approximately 350–450 proposals per year were contested by firms. The likelihood of exclusion during this time ranges from a high of 35% in 2005 to a low of 25% in 2015.

As discussed in Section 2, firms that seek to exclude proposals must provide reasons for the basis of seeking exclusion. Firms may seek exclusion of a particular proposal for multiple reasons. In responding to a firm's appeal to exclude a proposal, the SEC cites whether it agrees or disagrees with each reason cited by the firm. Managers have the ability to contest individual proposals for more than one reason, although the SEC only requires one accepted reason to permit a proposal's exclusion.

Table 2.2 details the reasons provided by firms, the frequency that reason is cited, and the

[10] Data on the shareholdings is less frequently available for proposals that are not contested by the firm and appear on the proxy statement. Out of a 100 randomly sample proxy proposals, data on shareholdings is only available for 59 of the shareholder proposals.

Table 2.1: *Shareholder Proposals Received, Contested, and Excluded by Meeting year*

Meeting year	# Shareholder Proposals	# Proposals Contested by Firms	# Proposals that SEC Allows Exclusion
2003	878	400	276
2004	956	430	282
2005	922	444	323
2006	905	378	270
2007	964	369	273
2008	1,138	433	350
2009	1,112	400	247
2010	1,014	377	284
2011	790	315	229
2012	897	334	259
2013	992	346	256
2014	943	302	227
2015	964	324	244
Total	12,475	4,852	3,520

This table shows the total number of shareholder proposals received by firms, the number of proposals contested by firms under Exchange Act Rule 14a-8, and the number that the SEC allows firms to exclude from their proxy by meeting year.

likelihood that the SEC permits an exclusion. The three most frequently cited reasons are when the proposal violates procedural requirements, includes false or misleading statements, or deals with ordinary business operations. Firms succeed in excluding proposals on the basis that the proposal violates procedural requirements or relates to ordinary business matters in 73% of cases. However, firms are considerably less successful in arguing that proposals include false or misleading statements as the SEC only grants an exclusion in 21% of these instances.

Shareholders submit a variety of different proposals. Table 2.3 describes the frequency of the different types of proposals and the likelihood that they are contested by firms. Social responsibility and environmental proposals are most commonly submitted by shareholders with 27% ($3,308$ out of $12,475$) of all proposals submitted by shareholders being of this type. Social/Environmental proposals also consist of the largest proportion of proposals that are contested by firms. Firms contest 40% ($1,331$ of $3,308$) of all social/environmental proposals they receive and the SEC offers the firm the opportunity to exclude in 69% (923 of $1,331$) of these cases. This is only

Table 2.2: *Frequency of Reasons for Exclusion of Shareholder Proposals*

Reason to Exclude Shareholder Proposal	Overall	SEC Does Not Allow Exclusion	% Overall	SEC Allows Exclusion	% Overall
Eligibility/Procedural requirements	1,481	401	27%	1080	73%
Includes materially false or misleading statements in proposal	1,148	909	79%	239	21%
Relates ordinary business operations	1,125	340	30%	785	70%
Already substantially implemented	706	377	53%	329	47%
Proposal violates state, federal, or foreign law	294	177	60%	117	40%
Company lacks power to implement proposal	278	221	79%	57	21%
Not subject for action by shareholders under the law of jurisdiction of firm	107	85	79%	22	21%
Related to Director Elections	126	57	45%	69	55%
Duplicative of another shareholder proposal	176	48	27%	128	73%
Conflicts with Company's Own Proposal submitted at the same meeting.	300	33	11%	267	89%
Related to resubmission of a proposal	80	6	8%	74	93%
Proposal is personal grievance or special interest	59	40	68%	19	32%
Relates to operations which account for < 5Relates to specific amount of dividends paid	31	5	16%	26	84%
Other	36	19	53%	17	47%
Total	5,986	2,750	46%	3,236	54%

This table provides the reasons that firms seek to exclude shareholder proposals under the Exchange Act Rule 14a-8 and the reasons the SEC allows the firm to exclude the proposal. The tables provides all reasons for all contested shareholder proposals from 2003 – 2015. Firms can contest individual proposals for multiple reasons which explains why the overall total number of reasons (N=5,986) exceeds the number of contested shareholder proposals (N=4,852).

Table 2.3: *Types of Shareholder Proposals Received, Contested, and Excluded.*

Proposal Type	# Shareholder Proposals	# Proposals Contested by Firms	# Proposals that SEC Allows Exclusion
Social/Environmental	3,308	1,331	923
Compensation Related	2,385	928	611
Board and Committee Proposals	2,615	667	483
Antitakeover Devices	1,995	625	441
Other Miscellaneous Issues	900	666	610
Issues Related to Annual Meeting	745	424	286
Voting Proposals	390	117	76
Auditor Related	137	94	90
Total	12,475	4,852	3,520

This table describes the number of shareholder proposals received by firms, the number of proposals contested by firms under Exchange Act Rule 14a-8, and the number that the SEC allows firms to exclude from their proxy by proposal submitter type.

slightly lower than the 72% ($3,049$ of $4,226$) likelihood that the SEC provides for exclusion of all contested shareholder proposals.

Table 2.4 displays descriptive statistics of firms that contest at least one proposal that they receive. Firms that contest proposals receive 3 proposals on average per year and contest nearly 2 on average per year. In comparison to firms that receive but do not contest any proposals, several differences emerge. In particular, firms that contest proposals tend to be larger (37.1 billion vs. 14.7 billion in market capitalization, t:stat: 15.9) and receive more proposals on average (3.1 vs. 1.7 proposals, t:stat: 17.8).

2.3.3 Who Receives and Contests Proposals?

We first examine the characteristics of firms that receive proposals. We do so by examining the probit model in Table 2.5, model (1):

Table 2.4: *Summary Statistics of Firms Contesting Proposals*

	Contested firms			Non-contested firms		
	N	Mean	Median	N	Mean	Median
Market Cap	2,426	37,146	13,076	2,563	14,763	5,124
Return	2,422	0.02	-0.00	2,562	0.04	-0.00
ROA	2,375	0.13	0.12	2,458	0.13	0.12
Leverage	2,417	0.27	0.25	2,563	0.27	0.26
% Inst. Holdings	2,379	0.69	0.72	2,502	0.72	0.76
Board Size	2,331	11.2	11.0	2,511	10.2	10.0
CEO is Chair	2,331	0.7	1.0	2,511	0.6	1.0
SP500	2,426	0.7	1.0	2,563	0.5	1.0
Total Comp	2,171	11,453	9,606	2,217	8,464	6,470
Litigation	2,426	0.06	0.00	2,563	0.04	0.00
Total Proposals	2,426	3.1	2.00	2,563	1.74	1.0
Total Contested	2,426	2.0	1.0	2,563	0.0	0.0

This table shows the summary statistics of firms with contested and non-contested proposals from the CRSP/COMPUSTAT population. The analysis is at a unique firm-year level in the period 2003–2015. Market Cap is share price times the number of shares outstanding as of reporting date and expressed in millions of dollars. Excess Returns is prior year's firm return minus the value weighted CRSP portfolio return. Leverage is the ratio of a firm's long term debt plus its current liabilities divided by assets. Accounting related variables are winsorized at 1%. ROA is defined as Income before Extra Items divided by the prior year's total assets. Pay Dividend is an indicator which equals one if the firm pays a dividend and zero otherwise. % Inst. Holdings is the percentage of shareholders held by institutional shareholders. CEO is Chair is an indicator which is equal to one if the CEO is also the Chairman of the board and zero otherwise. Total proposals is the total number of shareholder proposals that a firm receives in a year. Total contested is the number of shareholder proposals that a firm contests in a year.

$$\Pr(\text{Receive Proposal})_i = \alpha_i + \beta_1 \text{Mkt Cap}_i + \beta_2 \text{Return}_i + \beta_3 \text{ROA}_i + \beta_4 \text{Leverage}_i \\ + \beta_5 \text{\% Inst. Holdings}_i + \beta_6 \text{Board Size}_i + \beta_7 \text{CEO is Chair}_i + \beta_8 \text{Litigation}_i + \epsilon_i$$

The dependent variable *Receive Proposal* is an indicator variable that takes the value one when the firm receives at least one shareholder proposal in a given year and zero otherwise. Following the prior literature on shareholder proposals we include a variety of explanatory variables including size, firm performance, and governance (John and Klein, 1998; Karpoff *et al.*, 1996). Prior work has suggested that firms with larger boards and where the CEO is also the chairperson indicates lower quality governance (Yermack, 1996; Hallock, 1997). Therefore, we include variables measuring these two features in the regression to capture the quality of governance in the target firms.

Results are presented in Table 2.5, Panel A. Model 1 presents the baseline estimates. We find that firms that are larger, more leveraged, have greater institutional holdings, have larger board size, and where the CEO is also chairperson are all more likely to receive proposals from shareholders. Firms with worse accounting performance, as measured by ROA, or stock returns are also more likely to receive proposals in a given year. This result is consistent with shareholders being more interested at seeking changes at underperforming firms.

In model (2) of Table 2.5, we include a variable that identifies firms that receive a proposal in the prior year. The positive and significant coefficient on *Previously Received Proposal* suggests that firms are substantially more likely to receive a shareholder proposal in the current year if they received one in the year before suggesting that shareholders target the same firm for several years in a row. In particular, the coefficient values suggest that if a firm received a proposal in the prior year, it is 34% more likely to receive at least one in the current year. This variable is also economically significant in the sense that it contributes to a 14% increase in the adjusted-R^2 from 30% in model (1) to 44% in model (2).

In Models (3) and (4), we further explore by also including levels of expected and residual compensation.[11] Notably, firms with higher expected compensation are considerably more likely

[11] We define expected and residual compensation as in Core *et al.* (2008).

to receive a proposal in a given year. In particular, a one standard deviation increase in expected compensation from the mean increases the probability that a firm receives a proposal by 12% (i.e. from 8% to 20%). We do not find that a firm's level of residual compensation is associated with the propensity of receiving a proposal.

In Panel B, we explore the likelihood that firms receive specific types of proposals based on their actions. In models (1) and (2), we explore how varying level of executive compensation are associated with the firm receiving a compensation proposal (rather than any proposal as in Panel A). In these models, we find that greater levels of residual compensation are associated with the likelihood of receiving a shareholder proposal about compensation. In models (3) and (4), we explore whether firms that have committed environmental violations are more likely to receive proposals related to environmental and social issues. To find environmental violations, we utilize the Violation Tracker Database created by Good Jobs First that documents violations from the EPS, OSHA, the Mine Safety and Health Administration, and the Federal Railroad Administration beginning in 2010 that lead to penalties of $5,000 or more. We do find that firms that commit violations are statistically more likely to receive social proposals. In particular, the likelihood of receiving and environmental or social proposal increases by 3-fold for firms that have an environmental violation, however the propensity is still quite small in magnitude at 0.6% (those without a violation have a 0.2% probability in a given year).

In Panel C of Table 2.5, we examine whether firms that receive and contest proposals are more likely to be involved in proxy fights. We utilize FactSet's Shark Repellent database to find instances of proxy battles. In both instances, we find evidence that receiving and contesting proposals are associated with firms becoming engaged in proxy fights. In particular, firms that contest a proposal are 2% more likely (i.e., from a base probability of 0.3%) to become engaged in a proxy fight. The relatively low propensity reflects the relative infrequency of these events more broadly. This results suggests that shareholder proposals are part of a larger mosaic of shareholder intervention in companies and can often go together.

The set of contested proposals is a subset of the total set of submitted proposals. We examine the type of firms that contest proposals in more depth, and present the results from a probit model in Panel D of Table 2.5. In these regressions, the dependent variable equals one if the firm contests any of the proposals it receives and zero otherwise.

$$
\begin{aligned}
\text{Pr(Contest Proposal)}_i = {} & \alpha_i + \beta_1 \text{Mkt Cap}_i + \beta_2 \text{Return}_i + \beta_3 \text{ROA}_i + \beta_4 \text{Leverage}_i \\
& + \beta_5 \text{\% Inst. Holdings}_i + \beta_6 \text{Board Size}_i + \beta_7 \text{CEO is Chair}_i + \beta_8 \text{Litigation}_i \\
& + \beta_9 \text{Number Proposals} + \epsilon_i
\end{aligned}
$$

On the governance aspects, firms with larger boards and those with CEO-Chair duality are more likely to contest proposals. Following prior work that suggests that firms with larger boards and CEO-Chair duality have less effective governance mechanisms, this suggests that firms with less effective governance mechanisms are more likely to contest proposals. We also find that firms with greater institutional holdings are less likely to contest a proposal. Column (1) results suggest that a one standard deviation change in institutional ownership around the mean is associated with a 7% decrease in the likelihood of contesting a proposal.

Firms that receive more proposals are also more likely to contest at least one of the proposals they receive. In particular, economic magnitudes of the estimates suggest that firms that receive 3 proposals are 14% more likely to contest a proposal than firms that receive only 1 proposal (i.e., the likelihood of contesting a proposal rises from 43% to 57%). As with firms that receive proposals, firms that contest proposals in the prior year also are more likely to contest a proposal in a current year. Interpreting the magnitude of the coefficient on *Previously Contested Proposals*, if managers contest a proposal in the prior year, they are 25% more likely to contest a proposal in the current year. This suggests some stickiness in how managers at individual firms decide to address the proposals that they receive.

In models (3) and (4), we exclude firms that always contest proposals from our analysis. There is a subsample of firms that contest all proposals they receive (N=285) and therefore it may be firm policy to simply contest all proposals regardless of content or circumstances. When these firms are removed from the analysis, we continue to find similar results with the number of proposals the firm receives and whether they contested a proposal in the prior year being positively correlated with the choice to contest.

Overall, the results in Table 2.5 indicate both similarities and differences in the types of firms that receive proposals and those that seek exclusion of proposals. Larger firms are more likely to both receive proposals and contest proposals. In addition, firms that had previously received or contested a proposal in the prior year are also more likely to do so in the following year. On the other hand, firms with greater institutional holdings are more likely to receive shareholder proposals, but such firms are significantly less likely to contest the proposals they receive. This could arise from institutional holders using the proxy process as a medium for shareholder engagement.

2.3.4 Shareholding Levels of Proposal Submitters

We provide descriptive evidence on the extent of shares held by investors that submit proposals. As discussed earlier, shareholders have to own only $2,000 worth of shares or 1 percent of the share capital for one year to be able to file proposals. The low $2,000 requirement had led to calls for increasing the ownership levels required for shareholders to be able to file proposals.[12]

Table 2.6 provides descriptive statistics of the shareholdings data for proposals that are contested. The mean shareholding by investors that file contested proposals is $10.7 million and the median is $39,000. Univariate evidence suggests that larger shareholders have a greater success in seeing their proposals included in the proxy. The average shareholder whose proposal is contested and later settled has 20.1 million dollars in shareholdings on average as compared with 6.9 million on average for those that are contested and later excluded. This difference is statistically significant at the 1% level (t:stat 4.9).

[12] For instance, SEC Commissioner Daniel Gallagher argued that "activist investors and corporate gadflies have used these loose rules to hijack the shareholder proposal system," adding that "the stock ownership threshold for submitting shareholder proposals should increase from an 'absurdly low' $2,000 to as high as $2 million." (Remarks at the 26th Annual Corporate Law Institute, Tulane University Law School: Federal Preemption of State Corporate Governance Commissioner Daniel M. Gallagher New Orleans, LA March 27, 2014 (SEC, 2014))

Table 2.5: *Analysis of Firms that Receive and Contest Shareholder Proposals*

Panel A

	(1) Receive Proposal	(2) Receive Proposal	(3) Receive Proposal	(4) Receive Proposal
Mkt Cap(ln)	0.352***	0.237***	0.314***	0.243***
	(0.000)	(0.000)	(0.000)	(0.000)
Return	-0.197***	-0.125**	-0.227**	-0.154*
	(0.000)	(0.003)	(0.007)	(0.024)
ROA	-0.356**	-0.208*	-0.825***	-0.628***
	(0.001)	(0.030)	(0.000)	(0.000)
Leverage	0.434***	0.339***	0.646***	0.510***
	(0.000)	(0.000)	(0.000)	(0.000)
% Inst. Holdings	0.791***	0.536***	0.083	-0.059
	(0.000)	(0.000)	(0.603)	(0.681)
Board Size(ln)	0.581***	0.472***	0.366**	0.333***
	(0.000)	(0.000)	(0.003)	(0.000)
CEO is Chair	0.238***	0.203***	0.136**	0.157***
	(0.000)	(0.000)	(0.009)	(0.000)
Litigation	0.0818	0.119	0.0129	0.136
	(0.322)	(0.083)	(0.919)	(0.109)
Previously Received Proposal		1.644***		1.393***
		(0.000)		(0.000)
Expected (Total Comp)			0.075***	0.044***
			(0.000)	(0.000)
Residual (Total Comp)			0.001	0.003
			(0.742)	(0.239)
Constant	-5.936***	-4.931***	-4.809***	-4.247***
	(0.000)	(0.000)	(0.000)	(0.000)
Observations	51,727	49,846	20,695	19,570
R^2	0.30	0.44	0.27	0.40

Panel B

	(1) Comp Proposal	(2) Comp Proposal	(3) Env Social Proposal	(4) Env Social Proposal
Mkt Cap(ln)	0.354***	0.283***	0.478***	0.346***
	(0.000)	(0.000)	(0.000)	(0.000)
Return	-0.141	-0.080	-0.274**	-0.257**
	(0.071)	(0.248)	(0.003)	(0.010)
ROA	-1.255***	-1.143***	-0.394*	-0.284
	(0.000)	(0.000)	(0.037)	(0.132)
Leverage	0.704***	0.605***	0.492***	0.374**
	(0.000)	(0.000)	(0.000)	(0.006)
% Inst. Holdings	-0.117	-0.247	0.780***	0.396**
	(0.520)	(0.173)	(0.000)	(0.002)
Board Size(ln)	0.135	0.124	0.244	0.053
	(0.301)	(0.275)	(0.106)	(0.759)
CEO is Chair	0.102*	0.107*	0.300***	0.188***
	(0.039)	(0.027)	(0.000)	(0.000)
Litigation	0.071	0.150	0.083	0.018
	(0.525)	(0.132)	(0.518)	(0.884)
Expected (Total Comp)	0.042***	0.023*		
	(0.000)	(0.022)		
Residual (Total Comp)	0.009***	0.011***		
	(0.001)	(0.000)		
Committed Violation			0.352***	0.231**
			(0.000)	(0.001)
Previously Received Proposal		0.852***		1.232***
		(0.000)		(0.000)
Constant	-5.106***	-4.565***	-7.063***	-5.561***
	(0.000)	(0.000)	(0.000)	(0.000)
Observations	20,695	19,570	23,101	23,069
R^2	0.23	0.33	0.38	0.47

Panel C

	(1) Proxy Fight	(2) Proxy Fight
Receive Proposal	1.102***	
	(0.000)	
Mkt Cap(ln)	-0.195***	-0.133***
	(0.000)	(0.000)
Return	0.102**	0.077*
	(0.001)	(0.018)
ROA	-0.089	-0.103
	(0.457)	(0.331)
Leverage	0.006	0.098
	(0.964)	(0.430)
% Inst. Holdings	0.680***	0.701***
	(0.000)	(0.000)
Board Size(ln)	0.026	0.123
	(0.863)	(0.371)
CEO is Chair	0.018	0.037
	(0.775)	(0.527)
Contest Proposal		0.636***
		(0.000)
Constant	-2.110***	-2.544***
	(0.000)	(0.000)
Observations	51,727	51,727
R^2	0.11	0.04

Panel D

	(1) Contest Proposal	(2) Contest Proposal	(3) Contest Proposal	(4) Contest Proposal
Mkt Cap(ln)	0.116***	0.070***	0.217***	0.168***
	(0.000)	(0.000)	(0.000)	(0.000)
Return	-0.133*	-0.131	-0.190***	-0.191***
	(0.024)	(0.057)	(0.000)	(0.000)
ROA	-0.547**	-0.365*	-0.705**	-0.497*
	(0.007)	(0.048)	(0.004)	(0.031)
Leverage	0.074	0.085	0.228	0.237
	(0.639)	(0.577)	(0.144)	(0.120)
% Inst. Holdings	-0.483***	-0.451***	-0.022	0.025
	(0.000)	(0.001)	(0.888)	(0.871)
Board Size(ln)	0.425***	0.291***	0.551***	0.423***
	(0.000)	(0.000)	(0.000)	(0.000)
CEO is Chair	0.134*	0.114*	0.157*	0.127*
	(0.023)	(0.049)	(0.011)	(0.038)
Litigation	-0.027	-0.013	-0.047	-0.028
	(0.807)	(0.912)	(0.691)	(0.821)
Number Proposals	0.112***	0.087***	0.126***	0.102***
	(0.000)	(0.000)	(0.000)	(0.000)
Previously Contested Proposal		0.643***		0.635***
		(0.000)		(0.000)
Constant	-1.995***	-1.459***	-3.731***	-3.210***
	(0.000)	(0.000)	(0.000)	(0.000)
Exclude Always Contest Firms	No	No	Yes	Yes
Observations	4,593	4,266	4,267	3,959
R^2	0.09	0.11	0.15	0.17

This table presents results from probit regressions from the population of CRSP/Compustat firms that are merged with ISS Voting Analysis database. In Panel A, Regressions (1) and (2) have dependent variable equal to one if the firm receives at least one shareholder proposal that year and zero otherwise. Regressions (3) and (4) have a dependent variable equal to one if the firm (conditional on receiving a proposal) contests one of them and zero otherwise. In Panel B, the dependent variable is equal to one if the firm receives a compensation (models (1) and (2)) or environmental/social proposal (models (3) and (4)) in the subsequent year. In Panel C, the dependent variable is equal to 1 is the firm engages in a proxy battle in the subsequent year. In Panel D, the dependent variable is equal to 1 if the firm contests a proposal. See Table 2.4 for variable definitions. Standard errors (in parentheses) are double-clustered by firm and year. ***, **, * Indicate statistical significance at the 1%, 5% and 10% level respectively.

Table 2.6: *Share Holdings*

	N	Mean	Median	SD
All Contested	3,771	10,778	39	67,465
Contested and Voted	1,010	13,467	39	88,142
⇒ Pass	156	15,701	31	84,236
⇒ Fail	854	13,058	42	88,879
⇒ Mean diff.		2,643	t-stat: 0.3	
Contested and Not Voted	2,761	9,794	39	58,079
⇒ Withdrawn	607	20,153	133	90,436
⇒ Excluded	2,154	6,875	28	44,535
⇒ Mean diff.		13,278	t-stat: 4.9	

This table presents the dollar shareholdings of shareholder submitting proposals to firms that are contested by management. The data is manually collected from the sample of all proposals received and contested by management. All shareholdings are in thousands of dollars.

2.3.5 Contested Proposals that are Withdrawn

Among the $4,852$ proposals that firms contest from 2003–2015, some of these proposals will be placed on the proxy and voted on whereas others will be excluded from the proxy. The exclusion from the proxy can arise in several ways. The SEC could rule favorably for the firm and allows the proposal to be excluded from the proxy Alternatively, the firm can seek a dialogue with the shareholder to seek withdrawal of the proposal (i.e. the shareholder can then voluntarily remove the proposal from the proxy after it is satisfied with the firm's implementation of its suggestions). The firm can also simply decide to implement the proposal, thereby eliminating the need to place the proposal on the proxy.

We consider proposals as withdrawn when the shareholder formally withdraws the proposal in a written letter or the firm contacts the SEC to acknowledge that they substantially implemented the proposal (i.e. implicitly withdrawing the proposal since it is no longer a proposal for discussion). Out of the $3,632$ contested proposals which are never subject to a shareholder vote, 22% (i.e. 799 proposals) are withdrawn and never formally voted upon by shareholders.

In Table 2.7, we examine factors that contribute to the withdrawal of shareholder proposals. Firms that receive more proposals are less likely to withdraw a proposal. Moreover, firms with a

Table 2.7: *Analysis of Contested Proposals that Are Withdrawn.*

	(1) Withdrawn Proposal	(2) Withdrawn Proposal	(3) Withdrawn Proposal
Mkt Cap(ln)	0.133**	0.148**	0.123*
	(0.010)	(0.005)	(0.010)
Return	0.081	-0.001	-0.057
	(0.431)	(0.993)	(0.613)
ROA	0.084	0.153	-0.004
	(0.776)	(0.621)	(0.988)
Leverage	0.240	0.275*	0.318
	(0.057)	(0.038)	(0.064)
% Inst. Holdings	0.479*	0.495*	0.158
	(0.020)	(0.014)	(0.359)
Board Size(ln)	-0.378	-0.420*	-0.420*
	(0.063)	(0.041)	(0.026)
CEO is Chair	-0.126	-0.126	-0.092
	(0.182)	(0.197)	(0.320)
Shareholdings(ln)	0.084***	0.081***	0.033***
	(0.000)	(0.000)	(0.000)
Number Proposals	-0.039**	-0.039**	-0.036**
	(0.003)	(0.004)	(0.007)
Litigation	-0.163	-0.197	-0.196
	(0.195)	(0.125)	(0.193)
SEC Rejects Exclusion		1.378***	1.420***
		(0.001)	(0.001)
Constant	-2.122***	-2.231***	-1.365***
	(0.000)	(0.000)	(0.000)
Submitter Type	No	No	Yes
Proposal Type	No	No	Yes
Observations	2,283	2,283	2,283
R^2	0.06	0.10	0.20

This table presents results from regressions of sample proposals that are contested under Rule 14a-8 and later withdrawn. The probit model has a dependent variable equal to one if the proposal is withdrawn by the shareholder or substantially implemented and zero otherwise. Shareholdings describes the dollar value of shares held by the submitter of the proposal. SEC Rejects Exclusion is an indicator variable that takes on a value of one when the SEC does not allow the firm to exclude the proposal from its proxy. See Table 2.4 for other variable definitions. Proposal Type is an indicator variable for the subject matter of the proposal. Submitter type designates the type of shareholder submitting the proposal. Standard errors (in parentheses) are double-clustered by firm and year. ***,**,* indicate statistical significance at the 1%, 5% and 10% level respectively.

greater proportion of institutional shareholders are more likely to have a proposal withdrawn. A one standard deviation increase in institutional ownership around the mean is associated with a 4% increase in the likelihood of withdrawing a proposal.

We look at how variation in the amount of shares held by submitters influences the likelihood of a proposal being withdrawn. We hypothesize that shareholders with greater holdings are more influential and are more likely to have dialogue with managers that would facilitate implementation of their proposal prior to a shareholder vote. We found preliminary evidence of this in the descriptive statistics in Table 2.6. When included in the regression on Table 2.7, we find that the level of ownership is also significant. A one standard deviation increase in shareholder's stock holdings over the mean (i.e. from 10 million to 71 million) is associated with a 7% increase in the likelihood of withdrawing a proposal (i.e. from .39 to .46).

Managers can begin a dialogue with shareholders after contesting a proposal which could lead to a negotiated settlement prior to voting. The decision to withdraw can occur before or after the SEC reaches its decision of whether the proposal is allowed to be excluded from the proxy. In column (2) of Table 2.7, we examine how the SEC's rejection of managers' attempt to exclude the proposal (measured by indicator variable SEC Rejects Exclusion) influences the likelihood of its withdrawal. We find that once the SEC rejects the managers' attempt to exclude the proposal, the likelihood of having the proposal settled and withdrawn increases significantly. Specifically, when the SEC rejects managers' request to exclude the proposal, the likelihood of the proposal being settled increases by 50% (i.e. from .23 to .73). This suggests that the management becomes more willing to negotiate with shareholders once the SEC forces managers to include the proposal on their proxy (notwithstanding an alternative resolution with the shareholder).

Different types of shareholders and proposal types also potentially lead to variation in the likelihood of the proposal being settled and withdrawn. We classify shareholders who submit proposals into seven groups: hedge funds, socially responsible funds, pension funds, non-profits, individuals, groups (i.e. a mixture of different types of shareholders), and others.[13] We classify

[13] "Others" designates submitters that do not fall into any other easily distinguishable category with a significant number of shareholders. These include, for instance Loyola University (an educational institutional) and Marco Consulting Group (an investment consulting firm). Recent work that investigates activism by different constituencies

proposal types along eight different categories as described in Table 2.3. In Table 2.8, model 3 we add additional indicator variables for each of these submitter and proposal types. As compared with individuals (i.e. our base submitter category), proposals submitted by non-profits, socially responsible funds, and pensions are all more likely to be withdrawn (for the sake of parsimony we do not tabulate the submitter and proposal types). This is consistent with the impression that these groups are more willing to negotiate an amicable compromise with managers. As compared with the type of submitter, we do not find any particular category of proposals being more likely to be withdrawn. Although there is likely to be some selection on the types of proposals certain groups will submit, this suggests that the propensity of a proposal being withdrawn is related more highly to the type of shareholder submitting the proposal than the type of proposal itself. Including the submitter and proposal type improves the R^2 of the regression from 10 percent to 20 percent suggesting that these factors play a significant role in the likelihood of a proposal being settled.

2.3.6 Contested Proposals that are Placed on the Proxy and Voted

Shareholder proposals that a firm contests, but the SEC does not grant the right to exclude (or the firm does not reach some private resolution with the shareholder leading to its withdrawal) are placed on the proxy for a shareholder vote. Between 2003–2015, $1,209$ or 25% of all proposals that are contested by management appear on the proxy and are voted on by shareholders. Out of the $1,209$ proposals that are voted on by shareholders, 17% (209 proposals) ultimately gain shareholder approval. In Tables 8 - 10, we examine the factors that contribute towards their approval by shareholders.

We begin by in Table 2.8 examining all proposals that are voted on by shareholder to understand how contesting proposals and ISS support is associated with variation in the likelihood that a proposal ultimately passes once it is voted on. ISS is an influential proxy advisory firm that provides

includes pension funds (Gillan and Starks, 2000; Karpoff, 2001), hedge funds (Brav *et al.*, 2008; Klein and Zur, 2009), and social motivated groups (Agrawal, 2012).

recommendations to clients about how to vote in regards to shareholder proposals.[14]. The *ISS Support* variable is equal to one when ISS recommends supporting the proposal and management recommends voting against and zero otherwise.[15]

In Model (1) of Table 2.8, we find that contested proposals are unconditionally less likely to gain shareholder support. Specifically, contested proposals are 35% less likely to pass (from 36% to 1%). Proposals with ISS support, on the other hand, are substantially more likely (17%, from 6% to 23%) to pass. However, when we examine contested proposals that are also have ISS support, we find that the net effect is substantially. In particular, proposals that are contested, but that also have ISS support are 90% more likely to pass.

In Table 2.9 , we investigate the specific characteristics of those contested proposals that are approved by shareholders. Column (1) in Table 2.9 presents results from a probit regression where the dependent variable is one if the vote of the contested proposal is approved by shareholders and zero otherwise. Proposals at firms with greater extent of institutional investment are marginally more likely to gain majority support. When a proposal has ISS support, it is substantially more likely to receive majority support of shareholders. Specifically, the likelihood of the contested proposing passing is 20% higher if the proposal has ISS support and management rejects it. It should be noted that the decision for ISS to support a proposal endogenously reflects its opinion as well as those of its institutional clients.

In model (2) of Table 2.9, we include additional indicator variables for different submitter and proposal types. We find that the type of submitter has no impact on the likelihood that the contested proposal that is later voted upon gets approved. Notably, this differs from the results in Table 2.7 where the submitter type influenced the likelihood of withdrawal. We do find that social/environmental proposals and board proposals are significant less likely to receive shareholder approval.[16]

---

[14] A larger literature finds that ISS has a significant impact on voting (Cai *et al.*, 2009; Bethel and Gillan, 2002; Choi *et al.*, 2008; Alexander *et al.*, 2010; Daines *et al.*, 2010; Larcker *et al.*, 2015)

[15] If we simply examine the propensity for a proposal to pass if ISS supports the proposal (i.e. unconditional on management's view), it perfectly predicts whether a proposal passes. That is, when ISS does not support a proposal, it never passes during the sample period.

[16] This result differs from some prior research that suggests that proposals sponsored by institutional groups receive

Table 2.8: *Likelihood of Any Proposal Receiving Shareholder Approval*

	(1) Pass Proposal	(2) Pass Proposal	(3) Pass Proposal
Proposal Contested	-4.521***	-4.764***	-4.835***
	(0.000)	(0.000)	(0.000)
Contested = 1 AND ISS = For	4.416***	4.907***	4.916***
	(0.000)	(0.000)	(0.000)
ISS=For	0.813***	0.981***	0.791***
	(0.000)	(0.000)	(0.000)
Mkt Cap(ln)		-0.295***	-0.216***
		(0.000)	(0.000)
Return		-0.070	-0.063
		(0.280)	(0.457)
ROA		0.549	0.425
		(0.073)	(0.118)
Leverage		-0.471*	-0.454*
		(0.028)	(0.037)
% Inst. Holdings		0.911***	1.095***
		(0.000)	(0.000)
Board Size(ln)		0.233	0.157
		(0.084)	(0.330)
CEO is Chair		0.01	0.001
		(0.887)	(0.986)
Number Proposals		-0.030*	-0.020
		(0.043)	(0.156)
Constant	-1.248***	0.126	0.455
	(0.000)	(0.803)	(0.377)
Proposal Type	No	No	Yes
Observations	8,003	7,464	7,464
R^2	0.09	0.22	0.37

This table shows the likelihood of the passage of all shareholder proposals. The probit model has a dependent variable equal to one if the proposal is passed by the shareholder vote and zero otherwise. ISS Support is an indicator equal to one when ISS supports the proposal. Proposal Contested is an indicator equal to 1 if the proposal is contented by managers. Shareholdings describes the dollar value of shares held by the submitter of the proposal. See Table 2.4 for other variable definitions. All standard errors are double-clustered by firm and year. ***,**,* indicate statistical significance at the 1%, 5% and 10% level respectively.

Table 2.9: *Likelihood of Contested Proposals Receiving Shareholder Approval*

	(1) Pass Proposal	(2) Pass Proposal
Mkt Cap(ln)	-0.150**	-0.124*
	(0.002)	(0.021)
Return	0.231	0.081
	(0.192)	(0.604)
ROA	-0.329	-0.380
	(0.664)	(0.721)
Leverage	-0.473	-0.237
	(0.096)	(0.466)
% Inst. Holdings	1.035*	1.242*
	(0.021)	(0.027)
Board Size(ln)	-0.0797	-0.157
	(0.830)	(0.551)
CEO is Chair	0.161	-0.018
	(0.191)	(0.861)
ISS Support	1.468***	1.327***
	(0.000)	(0.000)
Shareholdings(ln)	-0.038	0.034
	(0.462)	(0.475)
Number Proposals	-0.037	-0.031
	(0.156)	(0.226)
Constant	-0.584	0.757
	(0.534)	(0.536)
Submitter Type	No	Yes
Proposal Type	No	Yes
Observations	957	950
R^2	0.24	0.43

This table shows the likelihood of the passage of shareholder proposals that were contested but later voted on. The probit model has a dependent variable equal to one if the proposal is passed by the shareholder vote and zero otherwise. ISS Support is an indicator equal to one when ISS supports the proposal and managers recommend voting against it. Shareholdings describes the dollar value of shares held by the submitter of the proposal. See Table 2.4 for other variable definitions. All standard errors are double-clustered by firm and year. ***,**,* indicate statistical significance at the 1%, 5% and 10% level respectively.

Most shareholder proposals require 50% approval by shareholders for passage (Cuñat *et al.*, 2012; Ertimur *et al.*, 2013).[17] In our sample, contested proposals that gain shareholder support win on average 65.3% of the vote, while those that do not pass receive an average approval rate of 22.1%. The average approval rate for all contested proposals is 29.5%. To put this number in comparison, the average approval rate for all shareholder proposals that are voted on is 36.1% suggesting that contested proposals that are voted on are only marginally less popular than shareholders proposals that are voted on in general.

In Table 2.10, we explore characteristics associated with explaining this variation in shareholder approval. The dependent variable is a continuous variable measuring the percentage of shareholder votes supporting the proposal. As suggested by Table 2.9, proposals that gain ISS support receive far higher shareholder approval. Also consistent with prior work (see Gillan and Starks, 2000), shareholder approval is higher when there are greater levels of institutional ownership. The level of shareholders held by the submitter is not significantly associated with the percent approval by shareholders overall (i.e. weakly statistically significant in regression (1), but not in (2)). All types of proposals, as compared to the base proposal type of antitakeover related device proposals, are less likely to gain as much approval.

Finally, in Table 2.11, we explore how receiving and contesting proposals is associated with major corporate changes including being acquired, CEO turnover, and a dividend increase. Investors who send shareholder proposals may be seeking broader changes in the firm than that specifically outlined in the proposal. With the greater public scrutiny that results from proposals, firms may be more willing to make such broader changes to the firm.

We find that receiving a proposal is associated with future changes in being acquired, CEO turnover, and dividend increases. Firms that receive a proposals are 4% more likely to have an increase in dividend payments the following year. We also find that firms that contest proposals are more likely to have CEO turnover and later acquisitions. In particular, firms that contest a

greater support (Gillan and Starks, 2000; Gordon and Pound, 1993; Thomas and Martin, 1998).

[17] In our sample, 1,019 of the 1,025 proposals that are contested and voted require 50% approval for passage. This is similar to the rate of all proposals that were not contested and voted in which 99% of these proposals required 50% approval to pass.

Table 2.10: *Level of Shareholder Approval for Contested Proposals*

	(1) Percent Approval	(2) Percent Approval
Mkt Cap(ln)	-0.013***	-0.005
	(0.000)	(0.070)
Return	0.041*	0.020
	(0.029)	(0.245)
ROA	-0.065	-0.046
	(0.359)	(0.458)
Leverage	-0.070*	-0.040
	(0.034)	(0.169)
% Inst. Holdings	0.155***	0.122***
	(0.000)	(0.001)
Board Size(ln)	0.013	0.000
	(0.783)	(0.994)
CEO is Chair	0.045**	0.027*
	(0.002)	(0.016)
ISS Support	0.307***	0.248***
	(0.000)	(0.000)
Shareholdings(ln)	-0.003	0.002
	(0.254)	(0.367)
Number Proposals	-0.000	-0.000
	(0.647)	(0.824)
Constant	0.135	0.349***
	(0.214)	(0.000)
Submitter Type	No	Yes
Proposal Type	No	Yes
Observations	955	955
R^2	0.566	0.718

This table shows the percentage of shareholder approval for proposals that were contested and later voted on. Percent approval is the percentage of shareholder votes supporting the proposal as designated in the firm's proxy statement. ISS Support is equal to one when ISS supports the proposal and managers recommend voting against it. Shareholdings describes the dollar value of shares held by the submitter of the proposal. See Table 2.4 for other variable definitions. All standard errors are double-clustered by firm and year. ***,**,* indicate statistical significance at the 1%, 5% and 10% level respectively.

proposal are 4% more likely to have a change in CEO in the next year respectively. These findings support that shareholder proposals appear to in some instances be "stalking horses" to achieving broader objectives that investors may desire.[18]

2.4 Discussion

Proposals that are contested by management and not allowed to be excluded from the proxy eventually reach a resolution by being placed on the proxy for a vote, implemented by the firm, or withdrawn by the shareholder (which is typically akin to being implemented by the firm). Successful shareholder proposals from the standpoint of the submitter can arise from either approval by shareholders or implementation by the firm.

Interestingly, we find a discrepancy in the types of contested proposals that are substantially implemented during negotiations and those that win as a result of shareholder votes. For example, over our sample period, 323 social and environmental proposals that were originally contested by firms are withdrawn by the shareholder or substantially implemented by the firm. However, virtually all of the 360 contested social/environmental proposals that are placed on the proxy for a shareholder vote win approval by shareholders. Managers may decide to implement the proposal to the submitter's satisfaction as a second best solution to avoid potential damage to a firm's public image in a visible corporate publicity battle. For instance, the activist group, the People for the Ethical Treatment of Animals (PETA), sent a shareholder proposal to Chipotle, a Mexican restaurant chain, calling on the firm to purchase chicken from suppliers that used certain less cruel slaughter methods. The management of Chipotle contested the proposal, but were not allowed to exclude the proposal. Rather than placing it on the proxy and potentially damaging the reputation of the brand, management substantially implemented PETA's proposal. According to PETA, "we purchase small amounts of stock; just enough to be able to submit shareholder resolutions . . . Our resolution called upon Chipotle to buy chicken from suppliers that used less

[18] We appreciate the reviewer for offering this idea for analysis.

Table 2.11: *Corporate Changes and Receiving/Contesting Proposals*

	(1) Acquired(t+1)	(2) Acquired(t+1)	(3) Is CEO Turnover	(4) Is CEO Turnover	(5) Dividend Increase	(6) Dividend Increase
Receive Proposal	0.091*		0.161***		0.106**	
	(0.010)		(0.000)		(0.005)	
Mkt Cap(ln)	-0.066***	-0.065***	-0.038***	-0.034***	0.166***	0.169***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Return	-0.071	-0.072	0.017	0.015	-0.087	-0.088
	(0.062)	(0.057)	(0.723)	(0.747)	(0.113)	(0.107)
ROA	0.238	0.239	-0.449***	-0.454***	1.006***	1.004***
	(0.057)	(0.054)	(0.000)	(0.000)	(0.000)	(0.000)
Leverage	0.265***	0.267***	0.128*	0.138*	0.051	0.054
	(0.000)	(0.000)	(0.044)	(0.026)	(0.434)	(0.412)
BTM	-0.000	-0.000	0.004	0.004	-0.016***	-0.017***
	(0.955)	(0.939)	(0.151)	(0.188)	(0.000)	(0.000)
% Inst. Holdings	0.228***	0.231***	0.020	0.022	-0.755***	-0.752***
	(0.000)	(0.000)	(0.665)	(0.634)	(0.000)	(0.000)
Board Size(ln)	0.295***	0.296***	0.155***	0.156***	0.920***	0.924***
	(0.000)	(0.000)	(0.001)	(0.001)	(0.000)	(0.000)
CEO is Chair	-0.032	-0.032	-0.190***	-0.189***	0.092***	0.094***
	(0.342)	(0.343)	(0.000)	(0.000)	(0.000)	(0.000)
Contest Proposal		0.125**		0.184***		0.090
		(0.003)		(0.000)		(0.053)
Constant	-2.676***	-2.682***	-1.229***	-1.253***	-3.071***	-3.093***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Observations	44,991	44,991	21,408	21,408	51,711	51,711
R^2	0.01	0.01	0.01	0.01	0.13	0.13

This table investigates how receiving a proposal and contesting proposals are related to firm acquisition, CEO turnover, and dividend increases in the following period. Acquired is defined as 1 if the firm is acquired during the following year. CEO Turnover is equal to 1 if the there is a change in the CEO during the following year as described in ExecuComp. Dividend increase is equal to 1 if there is a positive increase in the dividend payout between this year and the subsequent. See Table 2.4 for other variable definitions. All standard errors are double-clustered by firm and year. ***,**,* indicate statistical significance at the 1%, 5% and 10% level respectively.

cruel methods. They agreed to do just that in exchange for us withdrawing the resolution." [19]

We find that 17% of contested proposals win shareholder approval. Although this magnitude might appear low at first, it ought to be compared with the average level of support for non-contested proposals that are voted upon. Only 23% of non-contested proposals win shareholder support. While this is relatively higher, it suggests that contested proposals are not entirely frivolous claims by marginal shareholders.

The type of proposals that gain acceptance by shareholders is dynamic and changing over time. We classify the SEC's decision to exclude a proposal based on the exclusion criteria outlined in Table 2.2. However, over time the SEC has evolved in how it has interpreted these criteria. For example, in late 2002, the SEC decided to reinterpret equity compensation plans for senior executive directors as no longer ordinary business, but as matters of governance going forward. This change in interpretation created a significant change in the number of proposals that would previously have been excluded.[20] From 2003, the first year after the change in the SEC's interpretation, to 2015, 1,719 shareholder proposals were submitted and allowed to be placed on the proxy. Of these, 202, or 12%, won approval by shareholders. Without the change in the stance by the SEC, these proposals would have continued to have been excluded.

2.5 Conclusion

Proposals by shareholders offer a direct means for investors to seek changes at firms. However, not all proposals suggested by shareholders appear on the proxy. Following a set of criteria outlined in the Securities Exchange Act of 1934, managers can seek permission from the SEC to exclude certain types of proposals. In this paper we examine both the impetus for and consequences of contesting shareholder proposals. By investigating each stage in the process, we seek to understand how these proposals eventually come to a resolution by being dismissed by the SEC, implemented by the firm, or rejected by shareholders.

[19]"PETA's Shareholder Influence," QSR, Mark DeSorbo (accessed from `http://www2.qsrmagazine.com/articles/exclusives/0308/peta-1.phtml`)

[20]See "Division of Corporation Finance: Staff Legal Bulletin No. 14A" July 12, 2002 (SEC, 2002).

We show that managers often contest the proposals that they receive. From 2003–2015, 39% of all proposals suggested by shareholders are contested by management. However, in over a quarter of these cases, the SEC does not permit the firm to exclude the proposal. More significantly, the proposals that are contested by management represent more than the narrow interests of a minority shareholder. We find that 17% of proposals originally contested by managers, but placed to a vote, win shareholder approval.

Our analysis provides several opportunities for future inquiry. One important question is how the SEC decides to classify different matters and how this changes over time. While attorneys at the SEC may see a matter as part of ordinary business operations one year, the next they may see it as a governance matter. Such distinctions critically influence the type of matters that are potentially excluded from the proxy. Elucidating the process by which the SEC's interpretation evolves over time would provide insight into the types of matters that appear on the shareholder proxy statement.

We also find numerous proposals that are withdrawn by submitters after negotiations with the firm. The likelihood that managers appear willing to negotiate with the shareholder appears to rise after the SEC concludes that the firm cannot exclude the proposal from its proxy. Understanding how managers and shareholders undertake these private negotiations would provide deeper insights into how changes are made outside the proxy process.

Although we find that 17% of proposals that are contested by management are approved by shareholders, this does not mean they were necessarily implemented in full. Shareholder proposals are typically only advisory in nature.[21] By initially contesting the proposal, management conveyed their disinclination to implement the proposal. Consequently, it would be worthwhile to investigate whether proposals that are contested, but later approved by shareholders are more or less likely to be fully implemented by the firm.

A better understanding of contested shareholder proposals contributes to building a more

[21] In some states (e.g. Delaware), proposals can only be advisory in nature. As described by the SEC Division of Corporate Finance: "we have found that proposals that are binding on the company face a much greater likelihood of being improper under state law and, therefore, excludable under rule 14a-8(i)(1)." (Division of Corporation Finance, Staff Legal Bulletin No. 14) (SEC, 2001).

complete empirical picture of the governance process that occurs between shareholders and investors. The shareholder proposal process seeks to mediate conflict between the owners and managers of firms. To this extent that this process is influenced by managerial or regulatory (i.e. SEC) discretion, it can either improve the quality of the process or impede the resolution of shareholder–manager difference.

References

ACHARYA, V. V. and RYAN, S. G. (2016). Banks' financial reporting and financial system stability. *Journal of Accounting Research*, **54** (2), 277–340.

ADMATI, A. R. and PFLEIDERER, P. (2009). The "Wall Street Walk" and shareholder activism: Exit as a form of voice. *Review of Financial Studies*, **22** (7), 2645–2685.

AGRAWAL, A. K. (2012). Corporate governance objectives of labor union shareholders: Evidence from proxy voting. *Review of Financial Studies*, **25** (1), 187–226.

AHMED, A. S., TAKEDA, C. and THOMAS, S. (1999). Bank loan loss provisions: a reexamination of capital management, earnings management and signaling effects. *Journal of Accounting and Economics*, **28** (1), 1 – 25.

ALEXANDER, C. R., CHEN, M. A., SEPPI, D. J. and SPATT, C. S. (2010). Interim news and the role of proxy voting advice. *Review of Financial Studies*, **23** (12), 4419–4454.

ATHEY, S. and IMBENS, G. (2015). Lectures on machine learning. *NBER Summer Institute 2015 Methods Lectures*.

BALLA, E. and MCKENNA, A. B. (2009). Dynamic provisioning: A countercyclical tool for loan loss reserves. *FRB Richmond Economic Quarterly*, **95** (4), 383–418.

BARTH, M. E. (2007). Research, standard setting, and global financial reporting. *Foundations and Trends (R) in Accounting*, **1** (2), 71–165.

BARTKUS, C. E. (2016). Appealing No-Action Responses under Rule 14a-8: Informal Procedures of the SEC and the Availability of Meaningful Review. *Denver Law Review*, **93**, 199—-211.

BASKIN, D. L. (1992). SEC accounting and reporting positions: Allowances for loan-loss issues. *Banking Policy Report*, **18** (4), 11–15.

BEATTY, A. and LIAO, S. (2011). Do delays in expected loss recognition affect banks' willingness to lend? *Journal of Accounting and Economics*, **52** (1), 1–20.

BEATTY, A. L., KE, B. and PETRONI, K. R. (2002). Earnings management to avoid earnings declines across publicly and privately held banks. *The Accounting Review*, **77** (3), 547–570.

BEBCHUK, L. A. and COHEN, A. (2005). The costs of entrenched boards. *Journal of Financial Economics*, **78** (2), 409–433.

BECHT, M., BOLTON, P. and RÖELL, A. (2007). Corporate Law and Governance. *Handbook of Law and Economics*, **2**, 829–943.

BECK, P. J. and NARAYANAMOORTHY, G. S. (2013). Did the SEC impact banks' loan loss reserve policies and their informativeness? *Journal of Accounting and Economics*, **56** (2), 42–65.

BERNANKE, B. (2009). Financial reform to address systemic risk. *Speech at the Council on Foreign Relations*, **10**.

BETHEL, J. E. and GILLAN, S. L. (2002). The impact of the institutional and regulatory environment on shareholder voting. *Financial Management*, **31** (4), 29.

BRAV, A., JIANG, W., PARTNOY, F. and THOMAS, R. (2008). Hedge fund activism, corporate governance, and firm performance. *The Journal of Finance*, **63** (4), 1729–1775.

BROWN, J. R. (2016). Protecting shareholders from themselves: The SEC and restrictions on shareholder voting rights. *Working Paper*.

BUSHMAN, R. M. (2015). Transparency, accounting discretion and bank stability. *Economic Policy Review*, **22** (1).

— and WILLIAMS, C. D. (2012). Accounting discretion, loan loss provisioning, and discipline of banks' risk-taking. *Journal of Accounting and Economics*, **54** (1), 1 – 18.

— and — (2015). Delayed expected loss recognition and the risk profile of banks. *Journal of Accounting Research*, **53** (3), 511–553.

CAI, J., GARNER, J. L. and WALKLING, R. A. (2009). Electing directors. *The Journal of Finance*, **64** (5), 2389–2421.

CARLETON, W. T., NELSON, J. M. and WEISBACH, M. S. (1998). The influence of institutions on corporate governance through private negotiations: Evidence from TIAA-CREF. *The Journal of Finance*, **53** (4), 1335–1362.

CASEY, C. J. (1983). Prior probability disclosure and loan officers' judgments: some evidence of the impact. *Journal of Accounting Research*, pp. 300–307.

CHOI, S., FISCH, J. and KAHAN, M. (2008). Director elections and the influence of proxy advisors. *SSRN Electronic Journal*.

CHOI, S. H. D. (2006). It's getting hot in here: The SEC's regulation of climate change shareholder proposals under the ordinary business exception. In *Duke Environmental Law & Policy Forum*, vol. 17, pp. 165–192.

COLLINS, J. H., SHACKELFORD, D. A. and WAHLEN, J. M. (1995). Bank differences in the coordination of regulatory capital, earnings, and taxes. *Journal of Accounting Research*, pp. 263–291.

CORE, J. E., GUAY, W. and LARCKER, D. F. (2008). The power of the pen and executive compensation. *Journal of Financial Economics*, **88** (1), 1–25.

CUÑAT, V., GINE, M. and GUADALUPE, M. (2012). The vote is cast: The effect of corporate governance on shareholder value. *The Journal of Finance*, **67** (5), 1943–1977.

DAINES, R. M., GOW, I. D. and LARCKER, D. F. (2010). Rating the ratings: How good are commercial governance ratings? *Journal of Financial Economics*, **98** (3), 439–461.

DANG, T. V., GORTON, G., HOLMSTRÖM, B. and ORDONEZ, G. (2014). *Banks as Secret Keepers*. Tech. rep., National Bureau of Economic Research.

DAVENPORT, T. (2004). A switch to FASB keeps loan-loss issue on track. *American Banker*, **169** (125).

DECHOW, P., GE, W. and SCHRAND, C. (2010). Understanding earnings quality: A review of the proxies, their determinants and their consequences. *Journal of Accounting and Economics*, **50** (2), 344–401.

DELOITTE (2015). FASB Finalizing Credit Impairment Guidance. *Heads Up*, **22**.

DUGAN, J. (2009). Loan loss provisioning and pro-cyclicality. Available online at `http://www.occ.treas.gov/news-issuances/speeches/2009/pub-speech-2009-16.pdf`.

EINAV, L. and LEVIN, J. (2014). Economics in the age of big data. *Science*, **346** (6210).

ERTIMUR, Y., FERRI, F. and OESCH, D. (2013). Shareholder votes and proxy advisors: Evidence from say on pay. *Journal of Accounting Research*, **51** (5), 951–996.

FALEYE, O. (2007). Classified boards, firm value, and managerial entrenchment. *Journal of Financial Economics*, **83** (2), 501–529.

FASB (2016). Accounting Standards Update No. 2016–13, Financial Instruments – Credit Losses (Topic 326). *FASB in Focus*, pp. 1–4.

FDIC (2012). Community Banking by the Numbers.

— (2014). FDIC Compliance manual. Accessed: 2014-10-10.

FERNÁNDEZ, S., PAGÉS, J. M. and SAURINA, J. (2000). *Credit growth, problem loans and credit risk provisioning in Spain*. Banco de España.

FERRI, F. (2012). 'Low-Cost' Shareholder Activism: A Review of the Evidence. *Research Handbook on the Economics of Corporate Law*, p. 192.

— and SANDINO, T. (2009). The impact of shareholder activism on financial reporting and compensation: The case of employee stock options expensing. *The Accounting Review*, **84** (2), 433–466.

FILLAT, J. L. and MONTORIOL-GARRIGA, J. (2010). Addressing the pro-cyclicality of capital requirements with a dynamic loan loss provision system. *Federal Reserve Bank of Boston, Boston MA*.

FISCH, J. (2008). The transamerica case. *The Iconic Cases in Corporate, Edited by Jonathan R Macey*.

FSB (2009). Report of the financial stability forum on addressing procyclicality in the financial system. Financial Stability Forum.

FURLONG, F. and KNIGHT, Z. (2010). Loss provisions and bank charge-offs in the financial crisis: lesson learned. *FRBSF Economic Letter*, **2010**, 16.

G20 SUMMIT (2008). Declaration of the summit on financial markets and the world economy, G20, Washington, DC, 15 November.

GAO-US (2013a). Causes and consequences of recent bank failures. *GAO–13–71*.

— (2013b). Causes and consequences of recent community bank failures. *GAO–13–704T*.

GILLAN, S. and STARKS, L. (2000). Corporate governance proposals and shareholder activism: The role of institutional investors. *Journal of financial Economics*, **57** (2), 275–305.

GILLAN, S. L. and STARKS, L. T. (2007). The evolution of shareholder activism in the united states. *Journal of Applied Corporate Finance*, **19** (1), 55–73.

GOLDSTEIN, I. and SAPRA, H. (2014). Should banks' stress test results be disclosed? an analysis of the costs and benefits. *Foundations and Trends (R) in Finance*, **8** (1), 1–54.

GOMPERS, P., ISHII, J. and METRICK, A. (2003). Corporate governance and equity prices. *The Quarterly Journal of Economics*, **118** (1), 107.

GORDON, L. and POUND, J. (1993). Information, ownership structure, and shareholder voting: Evidence from shareholder-sponsored corporate governance proposals. *The Journal of Finance*, pp. 697–718.

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

HALLOCK, K. F. (1997). Reciprocally interlocking boards of directors and executive compensation. *Journal of financial and Quantitative Analysis*, **32** (03), 331–344.

HAMILTON, J. D. (1994). *Time series analysis*, vol. 2. Princeton University Press, Princeton, NJ.

HARRIS, T. S., KHAN, U. and NISSIM, D. (2017). The Expected Rate of Credit Losses on Banks' Loan Portfolios. *SSRN eLibrary*.

HASTIE, T., TIBSHIRANI, R. and FRIEDMAN, J. (2009). *The Elements of Statistical Learning (Springer Series in Statistics)*. Springer, 2nd edn.

JIANG, L., LEVINE, R. and LIN, C. (2016). Competition and bank opacity. *Review of Financial Studies*, **29** (7), 1911–1942.

JIMÉNEZ, G., ONGENA, S., PEYDRÓ, J.-L. and SAURINA SALAS, J. (2012). Macroprudential policy, countercyclical bank capital buffers and credit supply: Evidence from the Spanish dynamic provisioning experiments. *National Bank of Belgium Working Paper*, (231).

JOHN, K. and KLEIN, A. (1998). *Shareholder proposals and corporate governance*. NYU Stern Working Paper.

KARPOFF, J., MALATESTA, P. and WALKLING, R. (1996). Corporate governance and shareholder initiatives: Empirical evidence. *Journal of Financial Economics*, **42** (3), 365–395.

KARPOFF, J. M. (2001). The impact of shareholder activism on target companies: A survey of empirical findings. *Unpublished Manuscript, University of Washington*.

KIM, C. N. and MCLEOD JR, R. (1999). Expert, linear models, and nonlinear models of expert decision making in bankruptcy prediction: a lens model analysis. *Journal of Management Information Systems*, **16** (1), 189–206.

KLEIN, A. and ZUR, E. (2009). Entrepreneurial shareholder activism: Hedge funds and other private investors. *The Journal of Finance*, **64** (1), 187–229.

KLEINBERG, J., LUDWIG, J., MULLAINATHAN, S. and OBERMEYER, Z. (2015). Prediction policy problems. *American Economic Review*, **105** (5), 491–95.

LAEVEN, L. and MAJNONI, G. (2003). Loan loss provisioning and economic slowdowns: too much, too late? *Journal of Financial Intermediation*, **12** (2), 178–197.

LARCKER, D. F., MCCALL, A. L. and ORMAZABAL, G. (2015). Outsourcing shareholder voting to proxy advisory firms. *The Journal of Law and Economics*, **58** (1), 173–204.

LEMKE, T. P. (1987). The SEC no-action letter process. *The Business Lawyer*, pp. 1019–1044.

LEVIT, D. and MALENKO, N. (2011). Nonbinding voting for shareholder proposals. *The Journal of Finance*, **66** (5), 1579–1614.

LIBBY, R. (1976). Man versus model of man: Some conflicting evidence. *Organizational Behavior and Human Performance*, **16** (1), 1–12.

MENGLE, D. L. (1990). Market value accounting and the bank balance sheet. *Contemporary Economic Policy*, **8** (2), 82–94.

MIAN, A., RAO, K. and SUFI, A. (2013). Household balance sheets, consumption, and the economic slump. *The Quarterly Journal of Economics*.

— and SUFI, A. (2009). The consequences of mortgage credit expansion: Evidence from the us mortgage default crisis. *The Quarterly Journal of Economics*, **124** (4), 1449–1496.

— and — (2011). House prices, home equity-based borrowing, and the US household leverage crisis. *American Economic Review*, **101** (5), 2132–56.

MORRIS, S. and SHIN, H. S. (2002). Social value of public information. *The American Economic Review*, **92** (5), 1521–1534.

MULLAINATHAN, S. (2014). Big data and the inductive method to theory testing: A framework with applications. Available at: `http://live.wavecast.co/res-annual-conference/session/hahn-lecture-big-data-and-the-inductive-method-to-theory-testing-a-framework-with-applications/`.

NAGY, D. M. (1997). Judicial reliance on regulatory interpretations in SEC no-action letters: Current problems and a proposed framework. *Cornell L. Rev.*, **83**, 921.

OCC (1998). Allowance for loan and lease losses. *Comptroller's Handbook*.

— (2013). Remarks by Thomas J. Curry Comptroller of the Currency Before the AICPA Banking Conference Washington, D.C., September 16, 2013. *OCC Publication*.

PALMITER, A. R. (1993). The shareholder proposal rule: A failed experiment in merit regulation. *Alabama Law Review*, **45**, 879.

PARRINO, R., SIAS, R. W. and STARKS, L. T. (2003). Voting with their feet: institutional ownership changes around forced ceo turnover. *Journal of financial economics*, **68** (1), 3–46.

PÉREZ, D., FUMÁS, V. S. and SALAS, J. S. (2011). Do dynamic provisions reduce income smoothing using loan loss provisons? *Documentos de trabajo del Banco de España*, (18), 5–30.

POZEN, R. (1994). Institutional investors: The reluctant activists. *Harvard Business Review*, **72** (1), 140–149.

PRESCOTT, E. S. and WALTER, J. R. (2015). Did the financial reforms of the early 1990s fail? a comparison of bank failures and FDIC losses in the 1986-92 and 2007-13 periods. *Federal Reserve Bank of Richmond Working Paper 15-05*.

PREVOST, A. K. and RAO, R. P. (2000). Of what value are shareholder proposals sponsored by public pension funds. *The Journal of Business*, **73** (2), 177–204.

RATNOVSKI, L. (2013). Liquidity and transparency in bank risk management. *Journal of Financial Intermediation*, **22** (3), 422–439.

RENNEBOOG, L. and SZILAGYI, P. (2011). The role of shareholder proposals in corporate governance. *Journal of Corporate Finance*, **17** (1), 167–188.

ROCHET, J.-C. (1992). Capital requirements and the behaviour of commercial banks. *European Economic Review*, **36** (5), 1137–1170.

RYAN, P. (1988). Rule 14a-8, institutional shareholder proposals, and corporate democracy. *Ga. L. Rev.*, **23**, 97.

RYAN, S. G. (2007). *Financial instruments and institutions: Accounting and disclosure rules*. John Wiley & Sons.

— (2012). Financial reporting for financial instruments. *Foundations and Trends in Accounting*, **6** (3–4), 187–354.

— and KEELEY, J. H. (2013). Discussion of "Did the SEC impact banks' loan loss reserve policies and their informativeness?". *Journal of Accounting and Economics*, **56** (2), 66–78.

SAURINA, J. (2009a). Dynamic provisioning: The experience of Spain. *The World Bank Group, Washington DC*, 7.

— (2009b). Loan loss provisions in Spain: A working macroprudential tool. *Revista de Estabilidad Financiera*, **17**, 11–26.

SEC (1997a). Proposed Rule: Amendments to Rules on Shareholder Proposals. Available at: `https://www.sec.gov/rules/proposed/34-39093.htm`.

— (1997b). Remarks by Michael H. Sutton, Chief Accountant SEC, before the AICPA Banking Conference Washington, D.C., November 7, 1997. *Current Developments in Financial Reporting: Perspectives from the SEC*.

— (2001). Staff Legal Bulletin No. 14 (CF). Available at: `https://www.sec.gov/interps/legal/cfslb14.htm`.

— (2002). Staff Legal Bulletin No. 14A (CF). Available at: `https://www.sec.gov/interps/legal/cfslb14a.htm`.

— (2009). Staff Legal Bulletin No. 14E (CF). Available at: `https://www.sec.gov/interps/legal/cfslb14e.htm`.

— (2014). Remarks at the 26th Annual Corporate Law Institute, Tulane University Law School: Federal Preemption of State Corporate Governance, March 27, 2014. *SEC Publication*.

SHIVDASANI, A. and YERMACK, D. (1999). Ceo involvement in the selection of new board members: An empirical analysis. *The Journal of Finance*, **54** (5), 1829–1853.

STANTON, P. R. (1999). SEC reverses cracker barrel no-action letter. *Washington University Law Review*, **77** (3), 979–992.

STEEL, R. S. (2016). The underground rulification of the ordinary business operations exclusion. *Columbia Law Review*, **116** (6).

STRICKLAND, D., WILES, K. W. and ZENNER, M. (1996). A requiem for the usa is small shareholder monitoring effective? *Journal of Financial Economics*, **40** (2), 319–338.

SUBRAMANIAM, C. and WANG, F. (2009). Performance-oriented shareholder proposals and ceo compensation. *Working Paper, University of Texas*.

THOMAS, R. and COTTER, J. (2007). Shareholder proposals in the new millennium: Shareholder support, board response, and market reaction. *Journal of Corporate Finance*, **13** (2-3), 368–391.

— and MARTIN, K. (1998). Should labor be allowed to make shareholder proposals? *Washington Law Review*, **73** (1).

TIBSHIRANI, R. (1996). Regression shrinkage and selection via the lasso. *Journal of the Royal Statistical Society. Series B (Methodological)*, pp. 267–288.

WALL, L. D. and KOCH, T. W. (2000). Bank Loan-Loss Accounting: a Review of Theoretical and Empirical Evidence. *Economic Review*, (Q2), 1–20.

WALTER, J. R. (1991). Loan Loss Reserves. *Economic Review*, (Jul), 20–30.

WRIGHT, W. T. and WILLINGHAM, J. J. (1997). A computational model of loan loss judgments. *Auditing A Journal of Practice & Theory*, **16** (1), 99–113.

YERMACK, D. (1996). Higher market valuation of companies with a small board of directors. *Journal of Financial Economics*, **40** (2), 185 – 211.

— (2010). Shareholder Voting and Corporate Governance. *Annual Review of Financial Economics*, **2** (1), 103–125.

ZIMMER, I. (1980). A lens study of the prediction of corporate failure by bank loan officers. *Journal of Accounting Research*, pp. 629–636.

Appendix A

Appendix to Chapter 1

A.1 Current GAAP Rules

The two statements relevant to bank loan loss accounting are FAS 5 ("Accounting for Contingencies"), and FAS 114 ("Accounting by Creditors for Impairment of a Loan"). FAS 5 (paragraph 8) requires that both the following requirements be met to recognize a loss in the bank's financial statements:

a. Information available prior to issuance of the financial statements indicates that it is probable that an asset had been impaired or a liability had been incurred at the date of the financial statements. It is implicit in this condition that it must be probable that one or more future events will occur confirming the fact of the loss.

b. The amount of loss can be reasonably estimated.

This is commonly referred to as the incurred loss model. The allowance is to cover loan losses that are "probable" and "estimable" on the date of the evaluation.

A.2 Data Sources

A.2.1 Variable Definitions

A.3 Empirical Model Appendix

To estimate future losses $\mu_t = \mathbb{E}(CO_t|\mathscr{I}_{t-1})$, where $\mathscr{I}_{t-1}$ contains all information up to and including time $t-1$. For e.g., CO_{t-1} , $COrate_{t-1}$, and macro measures M at $t-1$. Let $\Delta M_{t-1} = M_{t-1} - M_{t-2}$.

To estimate μ_t, the bank uses historical loan information, and estimate the $COrate$. That is,

$$\mu_t = \text{loan balance}_{t-1} * \widehat{COrate}_t \tag{A.1}$$

The estimates for $\widehat{COrate}_t$ can be obtained by considering historical chargeoff rates, and adjusting for the change in the macro economic information and hence can recovered by estimating $\widehat{COrate}_t = COrate_{t-1} + \beta \Delta M_{t-1}$. To illustrate let M be unemployment at a county level. Substituting in Equation A.1,

$$\mathbb{E}(CO_{i,t}) = \text{loan balance * COrate}_{i,t-1} + \text{loan balance *}\Delta\text{unemprate}_t \tag{A.2}$$

Dividing by the loan balance at $t-1$,

$$\frac{\mathbb{E}(CO_{i,t})}{\text{loan balance}_{i,t-1}} = \text{COrate}_{i,t-1} + \Delta\text{unemprate}_{t-1}$$

$$\text{COrate}_{i,t} = \text{COrate}_{i,t-1} + \Delta\text{unemprate}_{t-1}$$

In my data, bank i operates only in county j. Rewriting to include the subscripts,

$$\text{COrate}_{ijt} = \text{COrate}_{it-1} + \Delta\text{unemprate}_{jt-1}$$

Estimating the above

$$\begin{aligned}\text{COrate}_{ijt} = \beta_1 * \text{COrate}_{it-1} + \beta_2 * \Delta\text{unemprate}_{jt-1} + \\ \beta_3 * \mathbf{X}_{it-1} + \beta_4 * \mathbf{M}_{jt-1} + u_{ijt}.\end{aligned} \tag{A.3}$$

where **X** are bank-level controls, and **M** are county-level controls.

The loan losses are driven by shocks to the county and the type of loans that the banks hold.

To estimate unobserved heterogeneity, I introduce for fixed effects for bank α_i and rewrite the equation as,

$$\begin{aligned} \text{COrate}_{ijt} = \beta_1 * \text{COrate}_{it-1} + \beta_2 * \Delta\text{unemprate}_{jt-1} + \\ \beta_3 * \mathbf{X}_{it-1} + \beta_4 * \mathbf{M}_{jt-1} + \alpha_i + u_{ijt}. \end{aligned} \tag{A.4}$$

This is a panel with about 2000 banks and with quarterly loan loss data. So, $N \gg T$ where T is from 2002Q1 to 2012Q4.

A.4 Supplementary Tables and Figures

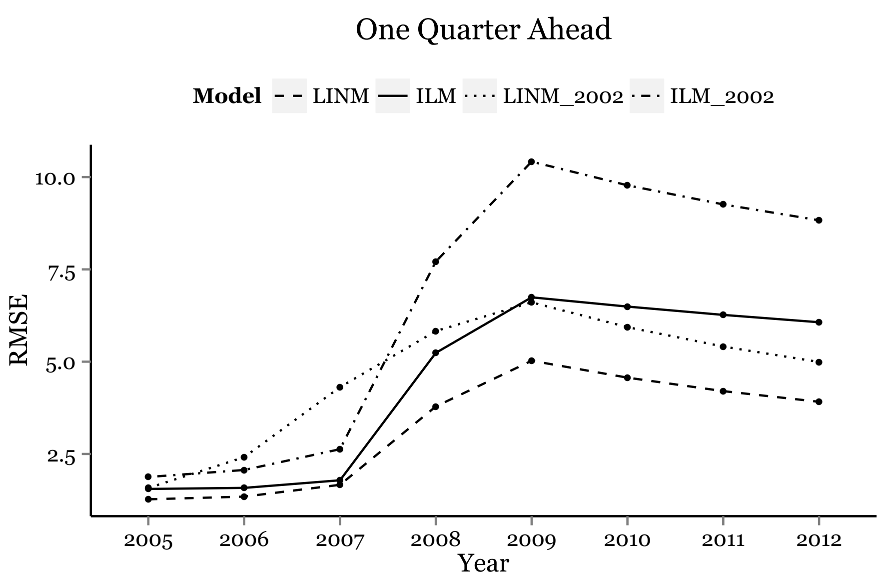


Figure A.1: *Allowance Model Cumulative Performance By Year.*

This figure presents evidence on the performance of the model predicted allowance relative to the current GAAP's ILM in their accuracy in predicting future losses. LINM is the root mean squared error, based on one-quarter out of sample, of the predicted allowance from the lasso model that takes the limited information as input. ILM is the root mean squared error, based on one-quarter out of sample, of allowance under current GAAP from the financial statements. LINM_2002 is the model estimated using the sample between 1996 – 2002 and ILM_2002 is the ILM model estimated using the sample between 1996 – 2002.

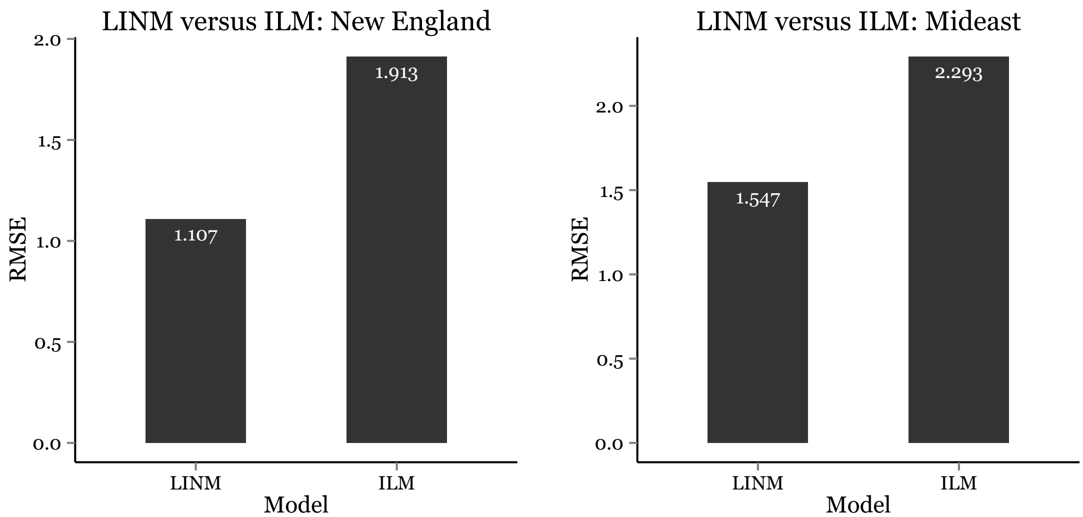


Figure A.2: *LINM vs ILM for county banks in BEA Region states*

This figure presents evidence on the difference in model performance based on sample banks that operate in particular states. The figures presents the performance of the LINM model predicted allowance relative to the current GAAP's ILM in their accuracy in predicting future losses. It shows the cumulative root mean squared error for the model estimated and tested out of sample in the period 2002–2012. LINM is the root mean squared error of the predicted allowance from the lasso model that takes the limited information as input. ILM is the root mean squared error of allowance under current GAAP from the financial statements. Each bargraph represents results from separate model estimations restricting to banks that operate in the corresponding BEA regions.

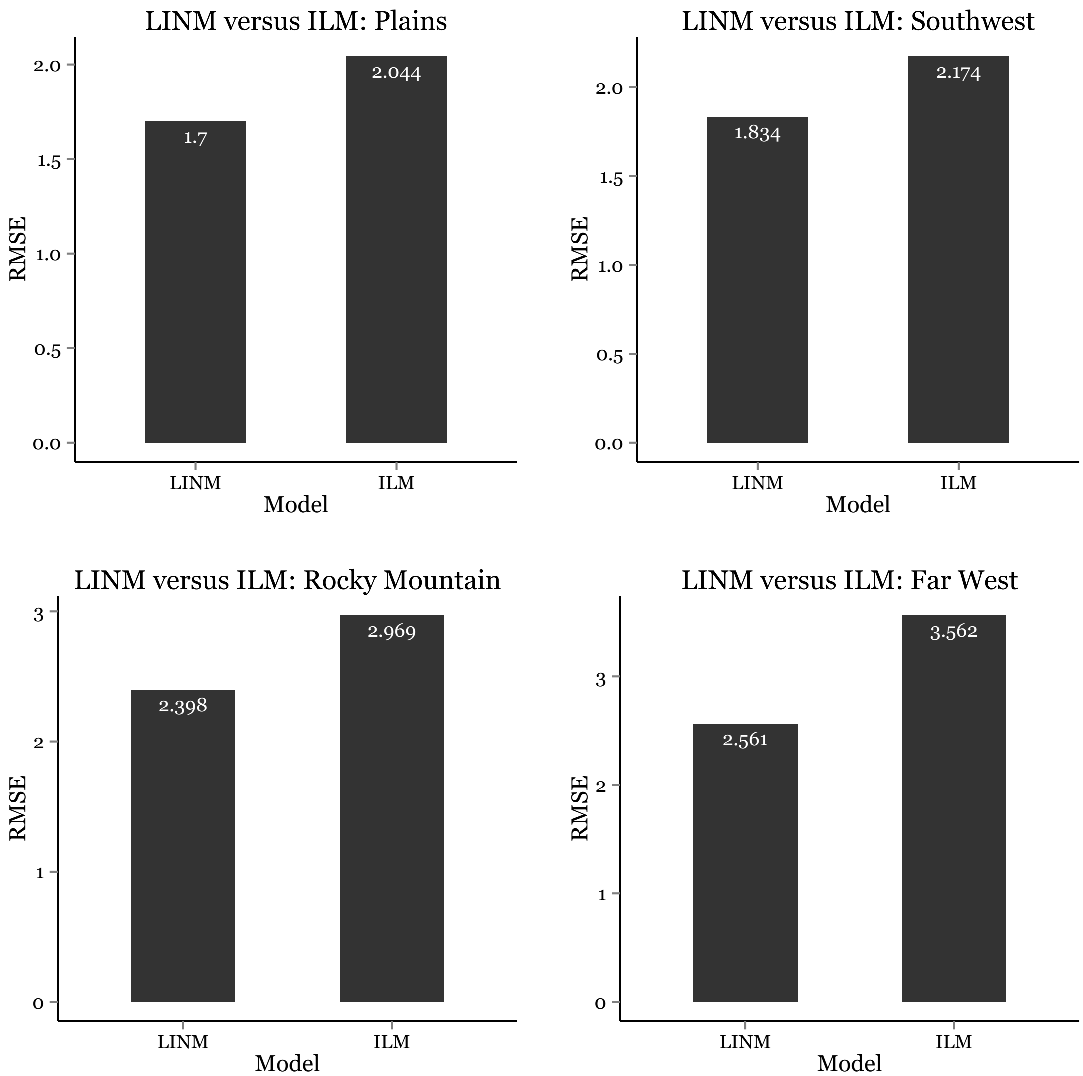


Figure A.3: *LINM vs ILM for county banks in BEA Region states*

This figure presents evidence on the difference in model performance based on sample banks that operate in particular states. The figures presents the performance of the LINM model predicted allowance relative to the current GAAP's ILM in their accuracy in predicting future losses. It shows the cumulative root mean squared error for the model estimated and tested out of sample in the period 2002–2012. LINM is the root mean squared error of the predicted allowance from the lasso model that takes the limited information as input. ILM is the root mean squared error of allowance under current GAAP from the financial statements. Each bargraph represents results from separate model estimations restricting to banks that operate in the corresponding BEA regions.

Table A.1: *Data Sources*

Data	Source
Bank balance sheet	FDIC Call Reports
Location information	FDIC Summary of Deposits
Home price growth	Corelogic/FHFA
Unemployment rate	Bureau of Labor Statistics
Income from tax returns	IRS
Business/Nonbusiness bankruptcies	Administrative Office of the U.S. Courts on behalf of the Federal Judiciary
Establishments - Number and Wages	QCEW/Bureau of Labor Statistics
Farm Income	Bureau of Economic Analysis
Per Capita Income	Bureau of Economic Analysis
County controls:	
Poverty rate	Census Bureau
High school	Census Bureau
Population demographics	Census Bureau

This table presents the list of data and their sources for variables used in the paper. All economic variables are identified at the county level.

Table A.2: *Variable Definitions*

Variable Name	Definition
Bank Variables	
Log Assets	Log of Total Assets
Total Loans	Total loans in banks portfolio
NCO	Net-chargeoffs
Four-qtr NCO + NPL	Sum of rolling four-quarter net-chargeoffs plus ninety-days past due and non-accrual loans at the end of the rolling window's fourth-quarter
Pct Four-qtr NCO + NPL	Four-qtr NCO + NPL scaled by total loans at the beginning of the quarter.
Chargeoffs	Amount that is charged-offs
Recoveries	Amount recovered in previously charged-off loans
Allowance	Loan Loss Allowance
Pct Allowance	Loan Loss Allowance scaled by total loans at the beginning of the quarter.
30 days PD	Loans that are 30-days past due scaled by total loans at the beginning of the quarter.
90 days PD	Loans that are 90-days past due scaled by total loans at the beginning of the quarter.
Nonaccrual	Loans that are in Nonaccrual scaled by total loans at the beginning of the quarter.
Delta 30 days PD	Change in loans 30 days past due divided by total loans at the beginning of the quarter.
Delta 90 days PD	Change in loans 90 days past due divided by total loans at the beginning of the quarter.
Delta Nonaccrual	Change in loans Non-accrual loans divided by total loans at the beginning of the quarter.
Pct RE Loans	Real estate loans as a percentage of total loans.
Pct CI Loans	Commercial and Industrial loans as a percentage of total loans.
Loans to Assets	Ratio of loans to assets
Securities to Assets	Ratio of the securities to assets
Interest Receivables	Income accrued but not yet collected on loans
Loan Yields	The ratio of tax-equivalent interest income divided by total loans
County Variables	
County Unemployment	County unemployment rate
HPI	Home price index
HPI Growth	Home price growth in the county
Establishments	Change in the number of business establishments in the county
Wage Growth	Change in adjusted gross income at the county
Farm Earnings	Change in farm earnings at the county
Per Capita	Per capita income at the county
Bus Bankr	Change in the ratio of business bankruptcies to all establishments in the county

This table presents the definitions for the variables used in the paper.

Table A.3: *Number of Sample Banks by Counties.*

Year	Mean	P25	Median	P75
1996	3.036	1	2	4
1997	2.911	1	2	3
1998	2.77	1	2	3
1999	2.679	1	2	3
2000	2.634	1	2	3
2001	2.601	1	2	3
2002	2.56	1	2	3
2003	2.495	1	2	3
2004	2.457	1	2	3
2005	2.409	1	2	3
2006	2.381	1	2	3
2007	2.339	1	2	3
2008	2.311	1	2	3
2009	2.287	1	2	3
2010	2.231	1	2	2
2011	2.162	1	1	2
2012	2.088	1	1	2

This table shows yearly distribution of number of sample banks by counties. For example, in 2002 the mean number of sample banks in a county is 2.56. The sample is formed by merging the FDIC summary of deposits data along with list of banks designated as community banks by the FDIC in the period.

Table A.4: *Coefficients from the LINM Lasso Model*

	1996 – 2002	1996 – 2005	1996 – 2006	1996 – 2007	1996 – 2008	1996 – 2010
(Intercept)	0.002	0.002	−0.004	−0.016	−0.021	0.001
30 days PD	0.077	0.090	0.089	0.095	0.129	0.167
90 days PD	0.183	0.268	0.304	0.324	0.355	0.388
Nonaccrual PD	0.223	0.274	0.291	0.318	0.406	0.522
30 days PD (t-3)	0	−0.011	−0.009	−0.006	−0.017	−0.023
90 days PD (t-3)	0	−0.034	−0.052	−0.058	−0.058	−0.064
Nonaccrual PD(t-3)	0.034	0.051	0.045	0.039	0.044	0.002
Loan to Asset	0	0.004	0.005	0.007	0.013	0.010
Securities to Asset	0	−0.000	−0.001	−0.001	−0.000	−0.005
Pct RE Loans	−0.007	−0.018	−0.018	−0.017	−0.014	−0.009
Pct CI Loans	0	−0.006	−0.008	−0.008	−0.009	−0.010
Log Assets	0.088	0.029	0.077	0.264	0.600	0.832
Net-Chargeoffs$_t$	0	0.133	0.157	0.160	0.211	0.334
Net-Chargeoffs$_{t-3}$	0.007	0.119	0.129	0.127	0.180	0.253
Net-Chargeoffs$_{t-6}$	0.006	0.105	0.112	0.110	0.108	0.152
Net-Chargeoffs$_{t-9}$	0	0.053	0.062	0.057	0.044	0.073

The table shows coefficients from the LINM model estimates for the sample banks. The LINM is the allowance prediction lasso model that takes the limited information as input. The coefficients of the variables are from lasso model estimated at bank-quarter level. Each columns represent coefficients from separate lasso models run between the periods 1996 – 2002, 1996 – 2005, 1996 – 2006, 1996 – 2007, 1996 – 2008 and 1996 – 2010 for sample banks. See Section A.2.1 for variable definitions.

Table A.5: *Implicit Coefficients from the ILM Model*

	Allowances					
	1996 – 2002	1996 – 2005	1996 – 2006	1996 – 2007	1996 – 2008	1996 – 2010
30 days PD	0.028***	0.030***	0.029***	0.028***	0.030***	0.032***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
90 days PD	0.013***	0.013***	0.017***	0.020***	0.024***	0.026***
	(0.002)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Nonaccrual PD	0.088***	0.102***	0.104***	0.104***	0.108***	0.126***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
30 days PD (t-3)	−0.013***	−0.017***	−0.017***	−0.016***	−0.016***	−0.016***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
90 days PD (t-3)	−0.007***	−0.007***	−0.010***	−0.012***	−0.015***	−0.016***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Nonaccrual PD (t-3)	−0.030***	−0.038***	−0.041***	−0.041***	−0.039***	−0.048***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Loan to Asset	−0.015***	−0.015***	−0.015***	−0.015***	−0.014***	−0.014***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Securities to Asset	0.002***	0.002***	0.002***	0.002***	0.002***	0.001***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Pct RE Loans	0.000	0.002***	0.001***	−0.000	−0.001***	−0.002***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Pct CI Loans	−0.003***	−0.003***	−0.004***	−0.004***	−0.005***	−0.006***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Log Assets	−0.156***	−0.149***	−0.150***	−0.147***	−0.152***	−0.138***
	(0.006)	(0.005)	(0.005)	(0.005)	(0.004)	(0.004)
Net-Chargeoffs$_t$	−0.033***	−0.049***	−0.046***	−0.044***	−0.026***	0.003
	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)
Net-Chargeoffs$_{t-3}$	−0.005	0.000	0.007*	0.013***	0.019***	0.057***
	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)
Net-Chargeoffs$_{t-6}$	0.053***	0.056***	0.058***	0.063***	0.071***	0.103***
	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)
Net-Chargeoffs$_{t-9}$	0.074***	0.075***	0.075***	0.080***	0.081***	0.109***
	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)
Observations	153,001	204,576	220,074	234,937	249,336	276,693
Adjusted R^2	0.837	0.798	0.781	0.763	0.748	0.718

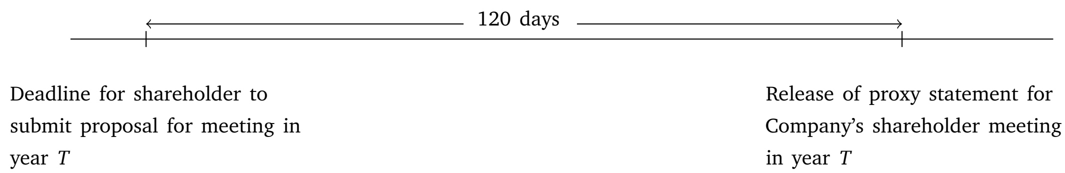
Notes: ***Significant at the 1 percent level.
**Significant at the 5 percent level.
*Significant at the 10 percent level.

This table shows OLS regressions to obtain coefficient of ILM for sample banks. The dependent variable is the allowances deflated by the beginning period loan balance. The coefficients are the implicit weights from the allowances recognized by banks. Each columns represent coefficients from separate regression models run between the periods 1996 – 2002, 1996 – 2005, 1996 – 2006, 1996 – 2007, 1996 – 2008 and 1996 – 2010 for sample banks. The regressions are at the bank-quarter level. All standard errors clustered at the county level. See Section A.2.1 for variable definitions.

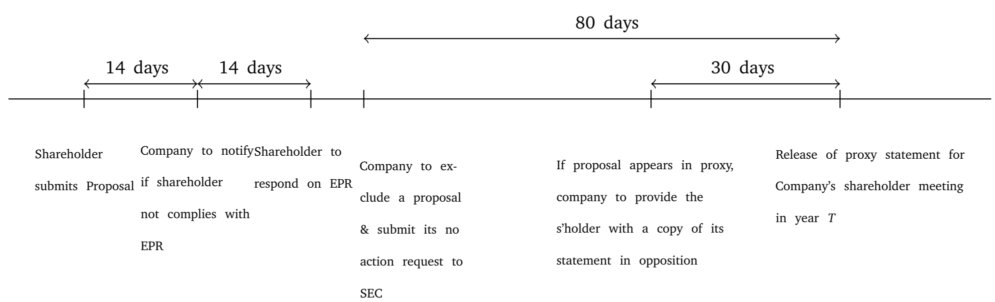
Appendix B

Appendix to Chapter 2

B.1 Timeline for Submitting Proposal



(a) *For Shareholder*



(b) *For Firm*

Figure B.1: *Timeline for Shareholder Proposal*

This appendix displays the timeline for filing and responding to shareholder proposals under Rule 14a-8 of the Securities Exchange Act of 1934. EPR is the "eligibility and procedural requirements" which requires the shareholder to hold $2,000 worth of shares or 1% of market value of equity continuously for at least a year.