Master's thesis

The association between prior earnings management and financial statements fraud

"Predicting financial statements fraud in the banking industry"

Erasmus University Rotterdam

Erasmus School of Economics

Department: Accounting, Auditing and Control

Author: Indra Tumbelaka (423098it)

Supervisor: Dr. Ferdinand Elfers

Co-reader: Drs. Rob van der Wal

Acknowledgements

I can finish this master thesis only because of the help of the people around me. First, I want to express my gratitude to my lovely wife, my parents, and all my friends. You all give me love more than I should receive. I also want to give my deep appreciation to my supervisor, Dr. Ferdinand Elfers who always give me valuable advice in my writing process, and to Drs. Rob van der Wal for the inputs. Next, I want to say thank you to Indonesia endowment fund for education (LPDP), Indonesia financial services authority (OJK), and central bank of Indonesia (BI) who give me the opportunity to continue my master study in Erasmus University Rotterdam. Finally, I hope this thesis is the start of my future academic writings. God bless us always.

Rotterdam, July 2017

Indra Tumbelaka

Abstract

How to predict financial statements fraud in the banking industry? To answer the question, this thesis examines the association between prior earnings management, the fraud incentives, and financial statements fraud in the banking industry. Using the abnormal loan loss provision and delayed loan loss recognition, this thesis finds new evidence that banks engaged with prior earnings management before committing financial statements fraud. Furthermore, this thesis finds that the cost of capital, minimum capital requirement, bank distress, and liquidity impact the likelihood of financial statements fraud. This thesis includes the financial crisis years in the sample period, therefore the time variable significantly impacts the results. Since the financial statements fraud and the abnormal loan loss provision increase in the financial crisis years, this thesis supports the implementation of expected loan loss provision (e.g. IFRS 9). Regulators, auditors, and other stakeholders can use the prediction models in this thesis to classify banks with a high risk of financial statements fraud.

Keywords: Financial statements fraud, accounting manipulation, earnings management, loan loss provision, delayed loan loss provision.

TABLE OF CONTENTS

TABLE OF TABLES

TABLE OF FIGURES

1. Introduction

The 2016 Global Fraud Study from Association of Certified Fraud Examiner (ACFE) reports the total loss of financial statements fraud reach $604 million. Banking and financial services industry report the most fraud cases and 12% of those cases are financial statements fraud (ACFE, 2016). Since banks' financial statements are used for both private and public decisions including monetary and financial stability policies, misleading information in banks' financial statements can cause damage not only to investors and creditors but also the economy.

This thesis examines the association between prior earnings management and the likelihood of financial statements fraud in the banking industry. Perols and Lougee (2011) find that prior earnings management can predict financial statements fraud. Dechow et al. (1996) mention that earnings management incentives can motivate firms to commit earnings manipulation. In the banking industry, Beatty and Liao (2014) find that earnings management through abnormal loan loss provision can predict the likelihood of provision manipulation. Therefore, the objective of this thesis is to answer the research question:

"Do prior earning management and fraud incentives increase the likelihood of financial statements fraud in the banking industry?

An answer of the research question is important since it can help stakeholders (e.g. regulators, investors, auditors) to predict and prevent financial statements fraud in the banking industry. For example, since Statements of Auditing Standards (SAS) require auditors to identify fraudulent financial reporting, this thesis can help auditors to identify banks with a high risk of accounting manipulation. In addition, understanding opportunistic earnings management and financial statements fraud behavior can also benefit in the effort to increase bank transparency (Bushman & Williams, 2015; Ma & Song, 2015).

This thesis uses financial restatements from auditors, banking regulators, and Securities and Exchange Commission (SEC) as the proxy of financial statements fraud. In order to only capture the fraud intention, this thesis excludes clerical and errors restatements. In addition to abnormal loan loss provision from Beatty and Liao (2014), this thesis uses delayed loan loss provision from Bushman and Williams (2015) to measure earnings management. The fraud incentives tested in this thesis are the cost of capital, minimum capital requirement, bank distress, and liquidity.

Following Dechow et al. (2011), and Beatty and Liao (2014), this thesis uses logit regression to observe the association between the explanatory variables and the financial statements fraud. In addition, inconsistent with Beneish (1999a) and Perols and Lougee (2011) this thesis uses random sampling of fraud and non-fraud observations to test the hypotheses. The observations of this thesis are the US bank holding companies from the year 2003-2015. Therefore, the sample period includes the years around the financial crisis. Following Beneish (1999a) and Dechow et al. (2011), this thesis uses the significant explanatory variables to develop fraud prediction models in the banking industry.

The results show that prior earnings management, and the fraud incentives significantly associated with financial statements fraud. First, it means that banks tend to engaged opportunistic accounting policy before committing financial statements fraud. Second, it shows that cost of capital and specific industry incentives which are the minimum capital requirement and bank distress motivate banks to commit financial statements fraud. These results consistent with Dechowt et al. (1996), Beneish (1999a), and Beatty and Liao (2014). This thesis also finds evidence that the fraud incentives do not increase nor decrease the prediction power of prior earnings management to financial statements fraud.

Furthermore, since this thesis includes financial crisis years in the sample period, the time variable significantly impacts the results. In the financial crisis years, the fraud frequency increase together with the abnormal loan loss provision. However, the positive and significant association between delayed loan loss recognition and financial statements fraud still holds in each year. This result is consistent with Bushman and Williams (2015) and Ma and Song (2016) who mention that delayed loan loss provision can capture bank opportunistic behavior and bank transparency. Moreover, the additional tests of bank distress and liquidity also support the main results.

This thesis uses the logit estimations to develop the prediction models. The models use prior earnings management, the fraud incentives, and bank profitability to predict financial statements fraud in the banking industry. This thesis compares two approaches from Beatty and Liao (2014) and Dechow et al. (2011) to classify the prediction models results. The results show that in the default classification[1], Beatty and Liao (2014) approach has higher accuracy of non-fraud prediction. However, Dechow et al. (2011) approach can predict fraud observations better

[1] Beatty and Liao (2014) classify banks with logit estimation more than 0.5 as manipulator banks. Dechow et al. (2011) classify firms with cut-off score more than 1 as firms with high probability of accounting misstatements.

than Beatty and Liao (2014) approach. Using Dechow et al. (2011) approach, the out of sample test results the type-1 error and type-2 error respectively 72.49% and 11.15%[2].

Contribution

This thesis contributes to two streams of financial accounting literature. First, the earnings management and the financial statements fraud. Second, the financial accounting in the banking industry. In the earnings manipulation literature, following Beatty and Liao (2014) future research suggestion, this thesis finds new evidence that together with the fraud incentives variables, prior abnormal loan loss provision significantly associated with financial statements fraud. The positive association between the fraud incentives and the likelihood of financial statements fraud, confirms the prior literature findings that fraud incentives can motivate firms to commit financial statements fraud in the banking industry (Dechow et al., 1996; Heally & Wahlen, 1999; Ronen & Yaari, 2008). In addition to Nichols, Wahlen, and Wieland (2009) who find the association between delayed loan loss recognition and bank conservatism, and Bushman and Williams (2015) who find the association between delayed loan loss recognition and bank transparency, this thesis extends the use of delayed loan loss recognition, and find new evidence that delayed loan loss recognition significantly associated with the likelihood of financial statements fraud. Different to the current financial statements fraud literature, this thesis includes the financial crisis period and the banking industry. Therefore, this thesis suggests that that the association between prior earnings management and financial statements fraud increase in the financial crisis years.

In the banking industry, this thesis contributes to the banking regulation study. Since this thesis suggests that abnormal loan loss provision and delayed loan loss recognition positively associated with financial statements fraud in the financial crisis years, this thesis in line with Bushman and Williams (2015) and Ma and Song (2016) who find there is an increase of bank opacity in the financial crisis years. However, inconsistent with Bushman and Williams (2012), this thesis supports the future implementation of expected loan loss provision by the banking and accounting authorities.

The ultimate aim of this thesis is to contribute to the financial statements fraud prediction model in the banking industry. The banking-specific fraud prediction models in this model can be

[2] The result is from the prediction model (a) i.e. using abnormal loan loss provision from Beatty and Liao (2014) model (a).

used by regulators and auditors as the first indicators of banks with a high risk of financial statements fraud.

Limitations

In line with Dechow et al. (2011) and Perols and Lougee (2011), this thesis based on the assumption that abnormal accrual can explain the likelihood of financial statements fraud. Therefore, the future researchers can examine the association between specific abnormal accrual and with specific financial statements fraud in the banking industry. Next, this thesis suggests that the relatively low abnormal loan loss provision several quarters before fraud period is because banks perform real earnings management through loans restructuring. The future research can examine the association between real earnings management and financial statements fraud in the banking industry. Next, because of the low fraud predictions, the prediction models in this thesis are less likely to predict fraud prediction. Despite following Dechow et al. (1996), Beneish (1999a) and Perols and Lougee (2011) who matched the fraud and the non-fraud observations. The future research can test the data analytics methods from Perols et al. (2016) to increase the accuracy of the prediction models. Finally, since the prediction models in this thesis use earnings management variable that calculated from future variables (i.e. next year non-performing assets), the models in this thesis cannot be used as ex-ante financial statements fraud prediction.

2. Theoretical background

This chapter contains the theory and the literature review sections. First, the theory section explains the background theory behind the hypotheses and the research design. Second, the literature review section discusses the relevant literature to this thesis.

2.1 Theory

This section elaborates the theory of financial statements fraud, earnings management, and banking regulation. It starts with the elaboration of the main interest in this thesis which is the financial statements fraud. Next, it discusses the earnings management theory and its incentives. Finally, it provides a brief background of financial accounting and supervision in the banking industry.

2.1.1 Financial statements fraud

2.1.1.1 Definition

From practitioners' perspective, Statements of Auditing Standards 99 (2002, p. 3) defines fraud as *"An intentional act that result in a material misstatement in financial statements that are the subject of an audit."* This definition mentions that fraud is an intentional action to misstate the financial reports, this intentional action separates fraud from error which is unintentional. Even though auditors do not responsible for determining whether an accounting misstatement intentional or unintentional, auditors still responsible for finding misstatement either it is fraud or unintentional error that can impact financial statements users' decisions.

More specific, Association of Certified Fraud Examiner (ACFE) in the report to the nation (2016, p. 90) defines financial statements fraud as *"A scheme in which an employee intentionally causes a misstatement or omission of material information in the organization's financial reports."* First, the term employee represents employees in general, including the executive or upper management. In addition, ACFE also mentions that financial statements fraud is most likely committed by a team of employees than by an employee. Second, this definition also uses the term intention to classify misstatements or omissions as financial statements fraud. However, since fraudulent managers are less likely to confess, it is difficult to observe the difference between unintentional error and intentional fraud.

In academic literature, researchers commonly use the term financial statements fraud interchangeably with accounting fraud (Ericson et al., 2006), fraudulent financial statements (Jones et al., 2008), earnings manipulation (Dechow et al, 1996; Beneish, 1999a), accounting misstatements (Dechow et al., 2011), and accounting irregularities (Price et al., 2011). The terms fraud and manipulation are more related with the negative intention of fraud. On the other side, the terms misstatements and irregularities tend to be more neutral and do not separate whether managers unintentionally or intentionally violate the standards. Since earnings information is more relevant to many financial statements users than the other information in the financial statements (Dechow et al., 1994), researchers use the terms earnings manipulation instead of accounting manipulation to stress that managers perform the manipulation to overstate or understate the earnings numbers. In this thesis uses, the term financial statements fraud is used interchangeable with the other terms.

This thesis uses financial statements fraud definition from Perols and Lougee (2011, p.40), *"financial statements fraud occurs when managers use accounting practices that do not conform to GAAP to alter financial reports to either mislead some stakeholders about the underlying economic performance of the company or to influence contractual outcomes that rely on reported accounting numbers."* First, following Dechow et al. (1996) and Beneish (1999a), this definition mentions that financial statements fraud violates the accounting standards (e.g. US GAAP, IFRS). Therefore, this definition clearly distinguishes earnings manipulation from earnings management which is committed within the accounting standard. It is essential to separate earnings manipulation and earnings management in this thesis, since this thesis uses earnings management as an indicator of financial statements fraud. Next, following Heally and Wahlen (1999) objective of earnings management, this definition also includes the objective of the accounting manipulation which is to alter the fair wealth transfer from the equity and debt investors (Stolowy & Breton, 2004). Therefore, this definition also in line with Picker et al. (2013) who mention that the objective of the fraudulent financial statements is to deceive or to influence financial statements users' decision about firm performance or condition.

2.1.1.2 Why it is important to study financial statements fraud

Based on the survey to public and private auditors around the world, including internal auditor, bank examiner, and computer forensic specialist, the total loss from fraud reached $6.3 billion and 9.6% of the total fraud cases reported are from financial statements fraud with the median loss almost $1 million or the biggest compared to the other two types of fraud (ACFE, 2016). However, because it is difficult to observe all the loss caused by the financial statements fraud, the loss amounts show in the report cannot entirely capture the actual total loss. In addition to the direct loss of financial statements fraud, the manipulation of accounting numbers can also cause more severe damage to the particular industry, such as the banking industry. The damages include loss of reputation and high-cost regulation and supervision.

In the agent and the principal relation, financial statements benefit to reduce the information asymmetry between the agent and the principle. However, since the agent has their own interest, and the principle has less information about the true company condition, the agent can use the information gap for their own benefit. Erickson et al. (2006, p. 113) mention that *"Some of the largest alleged accounting frauds in history occurred in the last several years, leading to the well-known upheaval in the accounting industry and sweeping legislative and regulatory changes. These events have left legislators, regulators, practitioners, and academics searching for answers about the causes of these alleged frauds. Understanding the underlying forces that gave rise to the alleged frauds is a necessary precursor to effectively preventing future occurrences. Many have suggested that the explanation lies in the incentives and opportunities for personal gain faced by executives."* In other words, they suggest that the key to preventing future financial statements fraud is the understanding of the managers' incentives to commit fraud. These incentives include managers' compensation and the other circumstances that can motivate managers to commit fraud.

Early financial statements fraud identification can increase industry credibility and protect stakeholders from the more severe damages. Therefore, the study of ex-ante fraud circumstances or the condition before the fraud event become more important. Two conditions that can be seen before the fraud events are the fraud incentives and the opportunistic earnings management behavior. Since the fraud triangle theory introduces the concept of the pressure which are incentives and motives, financial accounting researchers examine the association of financial statements fraud with its incentives. Dechow et al., (1996) examine the association between several manipulations incentives with the earnings manipulation. They find that a low cost of external fund and a debt covenant limitations can motivate earnings manipulation. In addition, Beneish (1999a) who uses several financial indicators to predict financial statements fraud, finds that managers' stock transactions strongly associated with earnings manipulation.

In the banking and financial intermediary industry, the information asymmetry issues become more critical. Therefore, after the financial crisis, regulators and researchers focus on the bank transparency policy, together with the effort to enhance the counter cyclical effect (e.g.

capital and liquidity. However, bank transparency has a negative association with earnings management (Bushman & Williams, 2012). Since earnings management positively associated with provision manipulation (Beatty & Liao, 2014), financial statements fraud can also decrease bank transparency.

2.1.1.3 Fraud prediction model

The positive accounting theory suggests that market unable to capture the true condition of the firms. Therefore, stakeholders use prediction model based on accounting numbers to evaluate firms' performance (Watts and Zimmerman, 1986). Together with restatements information, investors also use accounting misstatements prediction model for their short and long term decisions (Dechow et al., 2014). The two-well-known fraud prediction models are M-score and F-score respectively from Beneish (1999a) and Dechow et al. (2011). The researchers develop these two models by first examining the firms' characteristics differences between the fraud firms and the non-fraud firms. Next, Beneish (1999a) and Dechow et al. (2011) respectively use probit and logit regression to first examining the significant association between several variables and financial statements fraud, and second to develop their own fraud prediction model. Beneish (1999a) and Dechow et al. (2011) find that financial statements items and prior earnings management captured by abnormal accrual can predict financial statements fraud. Furthermore, the more recent study from Perols et al. (2016) suggest that data analytics methods can increase the fraud model prediction power.

In the banking industry, Beatty and Liao (2014) predict and find that current abnormal loan loss provision can provision manipulation captured by financial restatements and SEC comment letters. However, because they only use abnormal loan loss provision to predict the manipulation, the accuracy of their model relatively low. Following Dechow et al., (2011) and Beatty and Liao (2014), and unlike Beneish (1999a) and Perols and Lougee (2011) who oversample their fraud observations and undersample the non-fraud observations, this thesis uses random observations of fraud and non-fraud banks. First, this thesis analyzes the financial indicators differences between the two groups. Thereafter, this thesis examines the pattern of several financial indicators from periods before the fraud to one year after the fraud period. Next, using the logit regression, this thesis examines the association between the variables. Finally, this thesis uses the significant variables to develop fraud prediction models in the banking industry.

2.1.1.4 The proxy of financial statements fraud

Researchers commonly use Securities and Exchange Commission (SEC) Accounting and Auditing Enforcement Releases (AAERs) and accounting restatements as a proxy of accounting manipulation (e.g. Dechow et al., 1996; Beneish, 1999a; Perols & Lougee, 2011). SEC AAERs database based on the SEC investigation over the public firms with a strong indication of rules violation (e.g. accounting standards). Therefore, it provides a direct and solid proxy of accounting misstatements. However, because SEC AAERs usually comes from a severe accounting manipulation case, the frequency of the reports relatively rare, and less likely to capture relatively small accounting fraud.

Dechow et al. (2010) mention that together with SEC AAERs and internal control deficiencies, accounting restatements is a good proxy of earnings misstatements. Accounting restatements either from the regulators (e.g. banking or capital market regulators) or public accountants can capture both intentional and unintentional accounting fraud (Dechow et al., 2011). Following Beatty and Liao (2014), this thesis uses accounting restatements as the proxy of financial statements fraud. In addition to the financial restatements from the firms' auditors, the accounting restatements in this thesis also include restatements based on the SEC investigation. Furthermore, this thesis also uses restatements from the banking regulators.

2.1.2 Earnings management

2.1.2.1 Definition

According to Healy and Wahlen (1999, p. 368), *"Earnings management occurs when managers use judgment in financial reporting and in structuring transactions to alter financial reports to either mislead some stakeholders about the underlying economic performance of the company, or to influence contractual outcomes that depend on reported accounting numbers."* First, this definition implicitly mentions about managers' intention to use judgement. Managers use judgements and discretions when there is a room or a grey area in the standards or the regulations. For example, a current year discretion to write-off certain amount of loan loss provision. Furthermore, the direction and the type of the judgements based on the managers' objective, knowledge, and past experiences. For example, CFOs with prior auditor experience would use their past expertise in financial reporting or tax as one of their judgement considerations. Second, this definition mentions that managers can use discretion in accounting procedures, for example, financial assets classification or loan loss provision recognition in the banking industry. And real economic transaction, such as postponed production in the manufacturing industry or loan rescheduling in the banking industry. In the banking industry, Bushman (2014) mentions that the opportunistic accounting judgements can go hand in hand with real transactions decisions. Finally, this definition includes the aim of the earnings management which is to change the numbers in the financial statements in order to deceive financial statements users' decisions about the company condition, such as analyst forecasts or to influence the decisions that based on the financial statements information such as the debt covenants.

However, despite the negative reputation earnings management, managers can also use discretion in accounting practices for their stakeholders' benefit. Ronen and Yaari (2008, p. 26) define earnings management as follows: *"Earnings management is a collection of managerial decisions that result in not reporting the true short-term, value-maximizing earnings as known to management. Earnings management can be Beneficial: it signals long-term value; Pernicious: it conceals short- or long-term value; Neutral: it reveals the short-term true performance. The managed earnings result from taking production/investment actions before earnings are realized, or making accounting choices that affect the earnings numbers and their interpretation after the true earnings are realized."* In harmony with Healy and Wahlen (1999) definition, the first part of the definition mentions about managers decisions to alter the fair earnings numbers. The term short-term relates to the reports period as known by the managers. The second part of the definition separates the earnings management into its impact to the stakeholders, the positive, the negative, and the neutral impacts. This part clearly mentions that managers manage the accounting numbers to give more information that could benefit the stakeholders. For example, managers recognize more reserve as an anticipation of predicted economic slowdown in the short future. Finally, this definition also suggests that managers influence the reports numbers using accounting judgements and innate transactions.

Based on the definitions above, there are three aspects that can differentiate earnings management from financial statements fraud. First, in earnings management, managers can use their judgements to alter accounting numbers within the accounting standards (Dechow et al., 2000). On the other side, managers who commit earnings manipulation violates the accounting standards (Dechow et al., 1996; Beneish, 1999a). Next, even though it is difficult to observe, Dechow and Skinner (2000) stress that in manipulation circumstances, the managers show stronger intention to deceive than to perform earnings management, therefore it is a common assumption (e.g. Beneish, 1999; Dechow et al., 2011; Perols & Lougee, 2011) that managers involved in earnings management before committing financial statements fraud. Finally, earnings management can be performed in two ways, accounting judgment and real economic transaction, different to earnings manipulation that only focuses on the violation of the accounting standards.

2.1.2.2 Earnings management incentives

The positive accounting theory (Watts and Zimmerman, 1986) explains the reasons why financial statements users, including managers, accountants, and investors prefer to use particular accounting procedures and methods. The basic assumption in this theory is all financial statements users act to maximize their own benefit. For example, managers classify their stocks investments (e.g. as trading or available for sale) based on their positive impact on the earnings. Three common regularities investigated by the positive accounting theory are the bonus hypothesis, the debt/equity hypothesis, and the size or the political hypothesis.

Together with the agency theory, the positive accounting theory explains the reason behind managers' behavior. However, the specific opportunistic behavior to alter financial reports numbers must be studied together with its motives. Based on the positive accounting theory hypotheses, Heally and Wahlen (1999) suggest three main incentives that can drive earnings management behavior, they are the capital market expectations, the use of accounting in the contract, and the regulatory in20centives. In line with shareholders' objective, managers also aim to increase company market price. Since investors focus on the stock market price through the analyst prediction and analyst are more likely to use earnings numbers than cash flow numbers (Dechow, 1994), managers commonly influence the company market valuation through earnings overstatement or understatements. Researchers examine the capital market incentives during specific circumstances, such as when a company beating analyst forecast, initial public offerings, and merger or acquisition (Ronen & Yahri, 2008). The survey from Graham et al. (2005) and Dichev et al. (2013) suggest that managers believe the market does not like unpredicted events, therefore it is important to meet analyst forecast.

Managers commonly face several contracts that need to be fulfilled. Creditors use financial statements to evaluate companies' repayment capacity. Furthermore, predictive financial statements numbers are also used in the debt covenants between the companies and creditors (e.g. banks). For example, the debt covenants require the company to maintain certain numbers of cash and equivalent for the next couple periods. Moreover, since the violation of debt contract can lead to bankruptcy, financial distress can also motivate managers to manage their financial statements.

The other contract that also important is management compensation contract. Since the management compensation is based on the financial statements numbers, managers can also manage the earnings numbers to get higher compensation including equity based compensation (Armstrong et al. 2013).

Regulatory incentives such as the industry specific regulation, taxation, and political reasons (Ronen & Yaari, 2008) motivate managers to alter the financial statements numbers. First, in a specific industry such as in the banking industry, regulators require banks to maintain a minimum amount of capital or liquidity (Beatty and Liao, 2014). Ahmed et al. (1999) suggest that to avoid the minimum capital requirement violation, banks decrease their LLP, or understate their loan write-off. Next, managers can also understate earnings to avoid high tax expense (Ronen & Yaari, 2008). Lastly, political regulation and regulators scrutiny also motivates managers to manage their earnings numbers (Jones, 1991).

2.1.2.3 Measuring earnings management in the banking industry

In the banking industry, researchers focus on the loan related accounts (e.g. loan loss provision, charge-off, loan loss allowance) to study earnings management. In addition to the loan loss accounts, the other studies of earnings management focus on the asset sales and classification (e.g. Beatty & Liao, 2002; Bischof et al., 2016). However, since loan loss provision is the major accrual account in bank financial statements (Beatty & Liao, 2014), researchers focus on this account to study how manager opportunistic behavior in the banking industry. Different to the other accounts, the loan loss provision amounts in the financial statements do not base on the common transactions or contracts. First, the loan loss provision amount depends on the current and the future loan amounts. Second, it depends on the loan quality. Theoretically, the higher the loan quality, the lower the loan loss provision amount, and otherwise. Finally, because it depends on the debtor’s payment capacity, it is also affected by the macroeconomic conditions. In the end, there are many external and internal aspects that can affect the loan loss provision amount, and therefore managers usually use their judgement to determine the loan loss provision amount.

2.1.2.3.1 Loan loss provision model

Consistent with the accrual model in the general industry that developed from the Jones model, the loan loss provision model also distinguishes the total loan loss provision into the non-discretionary and discretionary accrual. The non-discretionary is the accrual part that cannot be managed by managers, and the discretionary is the accrual part that usually used to manage earnings. The early studies in loan loss provision behavior find that managers are more likely to reduce the loan loss provision in the high earnings period and otherwise (Ma, 1988; Beatty & Liao, 2002; Uygur, 2013). However, different to the common accrual model in the general industry, researchers use several different models to separate the non-discretionary and the discretionary or the abnormal loan loss provision. Beatty and Liao (2014) evaluate nine prior loan loss provision models and then generate their four model of loan loss provision as follow:

Equation 1

Model (a).

$$LLP_{j,t} = \alpha_0 + \alpha_1 \Delta NPA_{t+1} + \alpha_2 \Delta NPA_t + \alpha_3 \Delta NPA_{t-1} + \alpha_4 \Delta NPA_{t-2} + \alpha_5 \Delta SIZE_{t-1} + \alpha_6 \Delta Loan_t + \alpha_7 \Delta GDP_t + \alpha_8 \Delta CSRET_t + \alpha_8 \Delta UNEMO_t + \varepsilon_{j,t}$$

Model (b)

$$LLP_{j,t} = \alpha_0 + \alpha_1 \Delta NPA_{t+1} + \alpha_2 \Delta NPA_t + \alpha_3 \Delta NPA_{t-1} + \alpha_4 \Delta NPA_{t-2} + \alpha_5 \Delta SIZE_{t-1} + \alpha_6 \Delta Loan_t + \alpha_7 \Delta GDP_t + \alpha_8 \Delta CSRET_t + \alpha_8 \Delta UNEMO_t + \alpha_{10} \Delta ALW_{t-1} + \varepsilon_{j,t}$$

Model (c)

$$LLP_{j,t} = \alpha_0 + \alpha_1 \Delta NPA_{t+1} + \alpha_2 \Delta NPA_t + \alpha_3 \Delta NPA_{t-1} + \alpha_4 \Delta NPA_{t-2} + \alpha_5 \Delta SIZE_{t-1} + \alpha_6 \Delta Loan_t + \alpha_7 \Delta GDP_t + \alpha_8 \Delta CSRET_t + \alpha_8 \Delta UNEMO_t + \alpha_{10} \Delta CO_t + \varepsilon_{j,t}$$

Model (d)

$$LLP_{j,t} = \alpha_0 + \alpha_1 \Delta NPA_{t+1} + \alpha_2 \Delta NPA_t + \alpha_3 \Delta NPA_{t-1} + \alpha_4 \Delta NPA_{t-2} + \alpha_5 \Delta SIZE_{t-1} + \alpha_6 \Delta Loan_t + \alpha_7 \Delta GDP_t + \alpha_8 \Delta CSRET_t + \alpha_8 \Delta UNEMO_t + \alpha_{10} \Delta ALW_{t-1} + \alpha_{11} \Delta CO_t + \varepsilon_{j,t}$$

The dependent variable, LLP $_{j,t}$ is bank j loan loss provision in t time divided by bank j total loan in time t. Model (a) is the model that doesn't include both the charge-off (ΔCO) and the change of loan loss allowance (ΔALW) variables, model (d) is the model that includes both the variables and model (b) and (c) respectively includes only the ΔALW and the ΔCO variables. All the four models use the change of Non-Performing Assets (ΔNPA) before and after the period measured, this represented that banks already consider asset impairment in the future, and use past NPA to estimate the loan quality impairment. Moreover, the four models also considering the macro economic variables, which are the change on Gross Domestic Products (GDP), the change in unemployment rates (UNEMP), and the return on the Case-Shiller Real estate index (CSRET).

Beatty and Liao (2014) find that the error terms of each model can capture current abnormal loan loss provision. They also find abnormal loan loss provision has a positive and significant association with restatements and SEC comment letters. The logit regression used by Beatty and Liao (2014) as follows:

Equation 2

$$Restatement\ or\ comment\ letter = \alpha_0 + \alpha_1 ARES + \varepsilon_{j,t}$$

The restatement or comment letter variable is a dummy variable, 1 for banks with restatements or SEC comment letter, and 0 otherwise. ARES is the average absolute value of the four model residuals. In line with Beatty and Liao (2014), Ma and Song (2016) mention that abnormal loan loss provision can capture the earnings smoothing through loan loss provision. However, the magnitude of absolute loan loss provision can capture managers opportunistic better than the abnormal loan loss provision.

Another loan loss provision model is from Bushman and Williams (2012) who examine the association between earnings smoothing through loan loss provision and future looking of loan loss provision with banks risk-taking. Bushman and Williams (2012) model as follows:

Equation 3

$$LLP_{j,t} = \alpha_0 + \alpha_1 EBLLP_{j,t} + \alpha_2 \Delta NPA_{j,t+1} + \alpha_3 \Delta NPA_{j,t} + \alpha_4 \Delta NPA_{j,t-1} + \alpha_5 \Delta NPA_{j,t-2} + \alpha_6 T1CAP_{j,t-1} + \alpha_6 \Delta SIZE_{j,t-1} + \alpha_7 \Delta GDP_{j,t} + \varepsilon_{j,t}$$

Because Bushman and Williams (2012) model intentionally want to measure earnings smoothing using loan loss provision, they include earnings before loan loss provision ($EBLLP_{j,t}$) in their model. Bushman and Williams (2012) model also adds tier 1 capital divided by weighted assets ($T1CAP_{j,t-1}$) to capture the impact of the capital requirement to the loan loss provision. Regarding earnings smoothing, Bushman and Williams (2012) model is different to Beatty and Liao (2014) that focus to distinguish the abnormal loan loss provision to capture earnings management opportunistic behavior (Ma & Song, 2016).

2.1.2.3.2 Delayed loan loss recognition

In addition to the non-discretionary and discretionary loan loss provision models, Nichols et al. (2009) accommodate the time factor in loan loss provision recognition, and then they find that a more conservative bank recognizes more timely loan loss provision than a less conservative bank. Following Nichols et al. (2009) approach, recent studies starting to use the timely loan loss recognition or the delayed loan loss recognition together with the loan loss provision accrual model (e.g. Beatty & Liao, 2011; Bushman & Williams, 2015). Bushman and Williams (2015) compare the R-square of the following two models to measure the delayed loan loss recognition:

Equation 4

Model (1)

$$LLP_{j,t} = \alpha_0 + \alpha_1 \Delta NPA_{t-1} + \alpha_2 \Delta NPA_{t-2} + \alpha_3 \mathrm{Capital}_{t-1} + \alpha_4 \mathrm{EBLLP}_t + \alpha_5 \Delta SIZE_{t-1} + \varepsilon_{j,t}$$

Model (2)

$$LLP_{j,t} = \alpha_0 + \alpha_1 \Delta NPA_{t+1} + \alpha_2 \Delta NPA_t + \alpha_3 \Delta NPA_{t-1} + \alpha_4 \Delta NPA_{t-2} + \alpha_5 \mathrm{Capital}_{t-1} + \alpha_6 \mathrm{EBLLP}_t + +\alpha_7 \Delta SIZE_{t-1} + \varepsilon_{j,t}$$

Both models above accommodate the prior period change of non-performing assets ($\Delta NPA_{t-1}$and ΔNPA_{t-2}), one year lagged of tier one capital requirement ($\mathrm{Capital}_{t-1}$), earnings before loan loss provision (EBLLP_t), and the change of total assets ($\Delta SIZE_{t-1}$). However, since model (2) accommodates the change of the current and future non-performing assets (ΔNPA_t and ΔNPA_{t+1}), the higher the R-square difference between the model (2) over the model (1) represents the more timely the loan loss provision recognition (i.e. equation 10). Bushman and Williams (2015) then suggest that that less timely loan loss provision recognition reflects bank opportunistic behavior.

2.1.3 Regulatory background

2.1.3.1 Bank financial statements

The banking industry is a highly regulated industry because of the information asymmetry between banks with its shareholders and creditors (Beatty and Liao, 2014). In order to mitigate the information asymmetry, banks in the US need to prepare their financial statements based on the generally accepted accounting principle (GAAP) released by the financial accounting standards (FASB). In addition to the GAAP requirements, the banking industry is also required to meet specific regulation by the financial regulators. This specific regulation is intended to decrease moral hazard and adverse selection behaviors, and to increase bank ability to survive, especially in the crisis circumstances. Many countries have adopted Basel framework which requires banks to provide a minimum capital based on the weighted risk of the assets. To calculate banks capital adequacy ratio (CAR), banks and regulators use data from banks' financial statements. Furthermore, to increase banks transparency, the Basel II (i.e. pillar 3) requires banks to disclose their conditions to the market and banks' financial statements are the main sources of those disclosures (Bushman & Williams, 2012; Bushman, 2014).

Different to manufacturing or merchandising industries, monetary accounts dominate banks' balance sheet. These accounts such as loans and investments in the assets, and deposits on the other side. Because of the nature of the banks' financial statements, the banking industry is one of the industries that are impacted by the implementation of fair-value accounting (Beatty and Liao, 2014). In addition to fair value accounting, the other main topic in bank accounting is the loan loss provision behavior. The U.S. GAAP (i.e. statement of financial accounting No. 114 an amendment of FASB Statements No. 5 and 15 and Accounting Standards Update No.2010-20) requires banks to accrue losses from uncollectible loans. However, responding the financial crisis and to counter the pro-cyclicality effect, both FASB and international accounting standard board (IASB), in line with Basel Committee on Banking Supervision, have developed new accounting standard for loan loss provision. The future implementation of the new standard and IFRS 9 will require banks to switch from the current incurred loan loss provision to the expected loan loss provision (Bushman & Williams, 2012).

2.1.3.2 The banking regulators

Federal and states supervision agencies develop a banking supervision system in the US. The federal supervision agencies are Federal Reserve, Federal Deposit Insurance Corporation (FDIC), and Office of the Comptroller of the Currency (OCC). The state supervision agencies are the state financial supervision agency in each state that responsible to supervise state-chartered banks. Besides the banking dual-supervision system, the other major supervision agency is Security and Exchange Committee (SEC) who regulates the capital market and therefore also supervises listed banks. However, despite the multi supervision system, in addition to the GAAP financial statements, banks' financial reports are centered to federal financial institutions examination council (FFIEC) that requires uniform financial reports for the financial industry.

Banks in the US can be grouped based on their chartered, which are the state chartered banks and the national-chartered banks. State-charted banks are banks that only operated in the state where the bank is chartered and the banks supervised by the state financial supervision agency and mostly by the Federal Deposit Insurance Corporation (FDIC). However, if the state bank is a member of Federal Reserve, the bank is supervised by the Federal Reserve. National-chartered banks are supervised by the Office of the Comptroller of the Currency (OCC) and automatically a member of Federal Reserve. Since both types of banks products are directly used by personal, family, or household, state-chartered and national-chartered banks are supervised by the Office of the OCC. In addition, if the financial institution is not directly involved with the commercial customers and acts as a parent company, the institution is called the bank holding company (BHC) and supervise by Federal Reserve. Together with the National Credit Union Administration, and Consumer Financial Protection Bureau, the Federal Reserve, FDIC, and OCC compose the interagency body of FFIEC.

2.1.4 Summary of theoretical background

The study of financial statements fraud is important to predict and to prevent financial statements fraud in the future. In the banking industry, fraud cost becomes more severe because it can decrease bank transparency and then increase the bank systemic risk. This thesis uses the term financial statements fraud for accounting practices that violate the accounting standards. The definition distinguishes financial statements fraud from the earnings management which is committed within the accounting standards. In addition to the use of accounting judgement, managers can also use real-economic transactions to deceive financial statements users (Heally & Wahlen, 1999). For example, loans restructuring in the banking industry. The agency theory and positive accounting theory explain the agent, the principal, and the auditor role in the organization. Those theories also try to answer the incentives behind earnings management incentives, which are the bonus hypothesis, the contract hypothesis, and the size/political hypothesis (Watts & Zimmerman, 1978). Since financial statements fraud shares the same objective with the earnings management, and managers commonly perform earnings management before financial statements fraud, earnings management can predict potential financial statements fraud in the future. In the banking industry, researchers focus on the loan loss provision account to study earnings management. Since the determination of loan loss provision amount depends on manager's judgement, loan loss provision account captures the information asymmetry between the manager and the stakeholders. In addition to the general accounting standards and to the specific accounting standards (i.e. loan loss provisions), banks also need to maintain several requirements by the regulators, such as a minimum capital requirement and liquidity. These regulations create other incentives to perform earnings management or even financial statements fraud.

2.2 Literature review

Practitioners and researchers have continuously examined the association between several fraud incentives and financial statements fraud to develop a better fraud prediction model. Using the assumption that managers perform earnings management before they commit fraud, Dechow et al. (1996), Beneish (1999a) and Dechow et al. (2011) use earnings management characteristics to predict future financial statements fraud. In the banking industry, there is only limited literature that examines the association between financial statements fraud and earnings management, two of them are from Uygur (2013) and Beatty and Liao (2014). Regarding the use of delayed loan loss provision approach, this thesis also discusses Bushman and Williams (2015) study.

2.2.1 Financial statements fraud and the prediction model

2.2.1.1 Dechow et al. (1996)

This paper examines earnings manipulation and its association with several aspects before and after the earnings manipulation occurred. To examine the earnings manipulation, Dechow et al. (1996) compare the earnings manipulation firms with the match non-earnings manipulation firms. The earnings manipulation firms are firms from SEC AAERs between 1982 and 1992. These firms investigated by SEC due to GAAP violation, specifically in earnings manipulation.

Focus on the manipulation incentives, Dechow et al. (1996) examine the bonus and the debt hypotheses. In addition, they also consider the need of external funding and insider shares trading as the other earnings manipulation incentives. Using the logit regression, Dechow et al. (1996) results support the debt hypothesis. They find that earnings management firms have a positive and significant association with firms with a high demand of external finance. Next, regarding the firms' governance factor, Dechow et al. (1996) find that firms with weak governance indicators have a higher probability to commit earnings manipulation. For example, earnings manipulation firms are less likely have an outside blockholder.

In conclusion, Dechow et al. (1996) examine the SEC AAERs to distinguish earnings manipulation firms from non-earnings manipulation firms. In line with the following papers of accounting manipulation (e.g. Beneish, 1999a; Perols & Lougee, 2011), Dechow et al. (1996) assume that earnings manipulation firms also engaged earnings management. After comparing the test and the control groups, and then using the logit regression, Dechow et al. (1996) find that the bonus and the debt hypotheses positively associated with earnings management firms. Furthermore, they find that earnings manipulation is more likely associated with firms with weak governance indicators.

2.2.1.2 Beneish (1999a)

Beneish (1999a) M-score is one of the fraud prediction models which commonly used by practitioners (e.g. ACFA). In order to develop his model, first, he defines accounting manipulation as the accounting practices that violate accounting. This simple definition distinguishes accounting manipulation from earnings management. Based on the definition, he then classifies his sample into the manipulator firms and the non-manipulator firms (i.e. control group). Consistent with Dechow et al., 1996, Beneish (1999a) classifies firms with accounting restatements or investigated by the SEC as the manipulator firms. In addition, this paper also uses the news to track firms with accounting manipulation. Using data from year the 1982-1992, Beneish (1999a) uses and compares 74 manipulator firms from several industries to 2,332 non-manipulator firms. Beneish (1999a) intentionally uses more manipulator firms, because it is difficult to find many manipulator firms in the random sample. Perols et al. (2016) mention the difficulties to find manipulator firms as *"findings needles in a haystacks"*. These small numbers of the manipulator firms also impact the model ability to identify manipulator firms. However, to overcome the statistical weakness, Beneish (1999a) uses the weighted probit regression together with the unweighted probit regression.

Next, Beneish (1999a) develops several firm performance indicators to identify the firm condition in the manipulation year and before the manipulation year. The indicators are: Days' sales in receivables index (DSRI); Gross margin index (GMI); Asset quality index (AQI); Sales growth index (SGI); Depreciation index (DEPI); Sales, general and administrative expenses index (SGAI); Leverage index (LVGI); Total accrual to total assets (TATA); and Distribution of variables. Accommodating the time variable, Beneish (1999a) finds that before the manipulation year, the manipulator firms have higher growth and leverage than the non-manipulator firms. The last finding, in line with Dechow et al. (1996) findings that firms with high need of external fund are more likely to manipulate their earnings. However, it seems inconsistent with Beneish (1999b) who mentions that leverage doesn't motivate manipulation. Later, Dechow et al. (2011) mention that this inconsistency can be caused by the different sample used by the researchers.

Together with univariate analysis findings, the multivariate analysis findings confirm the prior literature findings that high accrual firms are more likely have low earnings quality or most likely involved in earnings management (Jones, 1991; Dechow, 1994). Finally, since the results of the weighted and the unweighted probit models show the similar results for all the indicators, and the R-square number of the unweighted model higher than the weighted model, Beneish (1999a) develops the M-score model based on the unweighted probit estimation model as follow:

Equation 5

$$M - Score = -4.84 + 0.92DSRI + 0.528GMI + 0.4404AQI + 0.892SGII + 0.115DEPI - 0.172SGAI + 4.679TATA - 0.327LVGI.$$

Since the M-score model can be used to classify the manipulator and non-manipulator firms, Beneish (1999a) admits the probability of the type-1 and the type-2 errors. Using investors' relative cost type-1 and type-1 errors of 20:1 and 30:1 respectively, the M-score model classifies a firm as a manipulator firm if the M-score is greater than -1.78. With that benchmark, the model misclassifies 26% and 13.8% the manipulator firms and non-manipulator firms respectively.

In conclusion, Beneish (1999a) develops a simple and efficient prediction model of financial statements fraud. The practitioners such as the ACFE members use the M-score model as the first red flag for the further investigation. This paper supports the Watts and Zimmerman (1978) theory that prediction model based on accounting number can give new information that cannot capture immediately by the market.

2.2.1.3 Dechow et al. (2011)

This paper follows Beneish (1999a) in elaborating firms' financial reports characteristics to develop a prediction model of accounting misstatement. However, this paper examines period from the year 1982-2005 or longer than Dechow et al. (1996) and Beneish (1999a). In the data collection, Dechow et al. (2011) mention several reasons why they prefer to use AAERs over accounting restatements or shareholder lawsuits data. Compare to accounting restatements, AAERs data are more likely to capture more severe accounting misstatement. Commonly SEC investigates firms with a high probability of fraud. However, consequently, AAERs data are less likely to capture the small magnitude of accounting misstatement or accounting misstatement that occurred in relatively small firms. On the other side, shareholder litigation can also capture significant manipulation. However, there is a possibility that the source of the lawsuit doesn't come from the accounting manipulation. In addition to the indicators of accounting misstatements in this paper, Dechow et al. (2010) suggest that SOX reports of internal control deficiency can also be an alternative misstatements indicator. This paper uses the term misstatement since the fraudsters usually deny the accusation of fraud.

Focus on firms' financial reports characteristics, this paper examines accrual quality, financial performance, non-financial measures, off-balance sheet activities, and market-based measures. In line with Perols and Lougee (2011), Dechow et al. (2011) find that prior to the misstatement year the accruals amount is higher or the same with the accruals amount in the misstatement year. Consistent with Beneish (1999a), this Dechow et al. (2011) mention that firms with larger assets are more likely to misstate their financial statements.

From the financial performance indicators, Dechow et al. (2011) find that misstatement firms tend to have decreasing returns on assets. Next, consistent with Dechow et al. (1996), this paper finds that the lack of liquidity over the last two years increase the misstatement probability. Finally, consistent with Dechow et al. (1996) and Perols and Lougee (2011), and inconsistent with Beneish 1999b, this paper finds that capital market incentives which are the need of capital and debt financing motivate firms to misstate their financial statements. In addition, compared to firms with higher leverage, firms with lower leverage have higher incentive to misstate their financial reports.

Dechow et al. (2011) also elaborates firms' accrual and financial characteristics prior to the manipulation year. They find that there are no significant differences between the accrual prior to the manipulation year with the accrual in the manipulation year. The possible explanation of finding is before the manipulation year the firm already manage the earnings within the accounting framework (i.e. earnings management). They also find that before the manipulation year there is a higher need for external funding.

The multivariate analysis in this paper uses logistic regression model that results in the F-score model as follow:

Equation 6

$$
\begin{aligned}
&Predicted\ Value/misstatements \\
&\quad = -7.893 + 0.790 \times (RSST\ accruals) + 2.518 \times (Change\ in\ receivables) \\
&\quad + 1.191 \times (Change\ in\ inventory) + 1.979 \times (\%Soft\ assets) \\
&\quad + 0.171 \times (Change\ in\ cash\ sales) + (-0.932) \times (Change\ in\ return\ on\ assets) \\
&\quad + 1.029 \times (Actual\ issue)
\end{aligned}
$$

The first variable, RSST accruals represents Richardson, Sloan, Soliman, and Tuna (2006) that captures working capital accruals together with long term assets and liabilities. Following the first variable, the second and the third variables respectively change in receivables and inventory also to capture short-term accruals. The fourth variable, soft-assets defines as assets outside cash and property, plant and equipment. The assumption is firm's short-term earnings target are more likely to achieve by firms with higher operation assets. The fifth and the sixth variables respectively change in cash sales and return on assets are indicators of firm's performance. The seventh variable represents the incentives from the market (Heally and Wahlen, 1999). The actual issue is a dummy variable to capture firm's new issuance of stocks or debts. Finally, using the real data test with the F-score greater than 1 means high possibility of accounting misstatements, F-score type-1, and type-1 errors respectively 36.31% and 31.38%.

In conclusion, Dechow et al. (2011) develop the F-score model to predict accounting misstatement. Together with M-score model from Beneish (1999a), the F-score model can help various stakeholders to capture firms with higher risk. Dechow et al. (2011) show that misstatement prediction model affects investors decisions. Different with Beneish (1999a), Dechow et al. (2011) F-score model use variables that can be taken directly from the financial statements. However, Dechow et al. (2011) develop the F-score model based on the general accrual assumption model, in fact, there are industries covered by this paper that use specific accrual model of earnings management such as the banking and the insurance industries.

2.2.2 Earnings management and opportunistic behavior in the banking industry

2.2.2.1 Beatty and Liao (2014)

This papers reviews and elaborates empirical literature of financial accounting in the banking industry. To the extent of earnings management in the banking industry, Beatty and Liao (2014) mention that unlike the accrual model in the non-bank industries, there are no dominant accrual models in the banking industry. Researchers in the banking industry do not develop their models based on one consensus model but use their own various assumptions to support their own model of abnormal accrual. In line with Beneish (2001), researchers in the banking industry examine capital and earnings management using specific accrual account, loan loss provisions. After the loan loss provision, have been excluded from tier-1 capital or in the new regime, Anandarajan et al. (2007) find that earnings management using loan loss provisions become more aggressive. Moreover, the earnings management through loan loss provision are higher in public banks than in private banks (Beatty and Liao, 2002). To the extent of capital management, Ahmed et al. (1999) find that the association between loan loss provisions and regulatory capital have decreased.

Next, Beatty and Liao (2014) evaluate nine different existing loan loss provisions models. They classify the models into three groups based on whether the models use the change of charge-off divided by total loans variable (ΔCO) or the change of the allowance variable (ΔALW). The first group contains the models with ΔCO (i.e. Beaver and Engel, 1996; Kim and Kross, 1998; Kanagaretnam et al. (2010);Beck & Narayanmoorth (2013). The second group contains the models without ΔCO but with ΔALW (i.e. Wahlen, 1994; Collins et al., 1995; Beat et al., 1995). Finally, the third group contains the models that do not include the two variables (i.e. Liu and Ryan, 2006; Bushman and Williams, 2012). Based on their evaluation of the three groups, Beatty and Liao (2014) then develop four models that represent all variables from the nine prior models. Further, using each model residual, Beatty and Liao (2014) perform logit regressions to examine if the abnormal loan loss provision can predict restatements and SEC comment letters related to loan loss provision. The results of the logit regression show that abnormal loan loss provisions can predict restatements and SEC comment letter.

In conclusion, Beatty and Liao (2014) provide new models of loan loss provision. These models can explain banks total loan loss provision better than the other prior models (e.g. Bushman & Williams, 2012). Using the loan loss provision models residuals as the proxy for earnings management, Beatty and Liao (2014) find a positive and significant association between financial statements fraud captured by restatements and SEC letter and earnings management. However, compare to prior accounting based fraud prediction model (e.g. M-score, F-score), the four logit regression models relatively provide low R-square and prediction power. The possible reason of the model weakness is the model only examine earnings management factor without considering earnings management incentives such as minimum capital and liquidity requirement.

2.2.2.2 Bushman and Williams (2015)

This paper uses bank data from the year 1993-2009 to examine the association of timely loan loss recognition with bank individual and systematic risk. Bank transparency has a significant association with bank risk (Bushman and Williams, 2012), therefore accounting policy through loan loss recognition can then affect bank risk through bank transparency. Since loan loss allowance represents a buffer to loan quality impairment in the future, delayed loan loss recognition captures opportunistic earnings management behavior by delaying current expense to the future period.

To measure delayed loan loss recognition, this paper uses the approach of Nichols et al. (2009), Beatty and Liao (2011) and Bushman and Williams (2012) by comparing R-square of loan loss provision accrual models. The first model is the model without considering the future non-performing assets, and the second model is the model with non-performing assets variable. However, different to Beatty and Liao (2011), Bushman and Williams (2012; 2015) accommodate bank size in their loan loss provision model. Regarding the impact of delayed loan loss provision to bank transparency, since bank financial statements provide significant information to stakeholders (Bushman & Williams, 2012), accounting choices have important roles in bank transparency. However, Bushman and Williams (2015) also admits, despite loan loss provision represents a major accrual in bank financial statements (Beatty & Liao, 2014), less information related to loan loss provision recognition cannot entirely represent the complexity of bank opacity.

Bushman and Williams (2015) find that delayed loan loss recognition has a positive and significant association with bank liquidity risk. Regarding the cost of capital, in line with Dechow et al. (1996), another inference of this result is a bank with high delayed loan loss recognition has a high cost of equity. In the other words, a bank with higher opportunistic behavior is more likely to have a higher cost of capital. Next, regarding the minimum capital requirement, this paper finds that high delayed loan loss recognition increase the possibility of future capital insufficiency. This result in line with Beatty and Liao (2011) who suggest that bank with higher delayed on loan loss recognition has a higher decrease in bank loan to meet the minimum capital requirement. Finally, this paper also finds that during the financial crisis, delayed loan loss recognition has a higher impact on bank negative equity returns.

In conclusion, Bushman and Williams (2015) provide new evidence that delayed loan loss recognition can capture management opportunistic behavior. They find that banks with higher delay on loan loss recognition tend to have higher liquidity risk and higher cost of capital.

2.2.3 Filling the literature gap

Prior literature in earnings management and manipulation provide evidence of the positive accounting theory hypotheses (Watts & Zimmerman, 1986: Heally & Wahlen, 1999) and the existing of opportunity manager behavior using accounting number (Dechow et al. 1996; Beneish, 1999a, Beneish 1999b; Dechow et al. 2011). Based on the association between earnings management, fraud incentives, and financial statements fraud, researchers develop fraud prediction models to help stakeholders classify firms with high risk of accounting manipulation. However, the prediction models commonly use general industries variables and therefore it is difficult to be applied in the banking industry. Therefore, the future research can focus on the association between fraud incentives and financial statements fraud in the banking industry.

In the banking industry, without examining the fraud incentives, Beatty and Liao (2014) show that earnings management positively associated with financial statements fraud captured by restatements and SEC comment letters. Focusing in the banking industry, this thesis aims to contribute to the accounting manipulation literature by examining fraud in one picture with earnings management and its specific incentives in the banking industry such as minimum capital and liquidity requirements. Accommodating the time factor, this thesis also examines ex-ante earnings management and its relation to the future financial statements fraud. In addition, using the timely loan loss provision recognition (Nichols et al., 2009; Bushman & Williams, 2012), this thesis also extends the use of timely loan loss provision recognition and test its relation with financial statements fraud.

2.2.4 Summary of literature review

Dechow et al. (1996) find that earnings management incentives can motivate earnings manipulations. Using firms' specific characteristics including earnings management and its incentives, Beneish (1999a) and Dechow et al. (2011) develop models based on the accounting numbers that can predict earnings manipulation or accounting misstatement. In the banking industry, Beatty and Liao (2014) provides evidence of the association between earnings management through loan loss provision and financial statements fraud captured by restatements and SEC. However, in their examination, Beatty and Liao (2014) do not consider earnings management incentives that probably increase the prediction power of their model. Using delayed loan loss recognition as the proxy of earnings management, Bushman and Williams (2015) find that bank with a higher delay on loan loss recognition more associated with bank opacity.

The current literature of the accounting manipulation studies the association of fraud incentives with the financial statements fraud. This is important to develop models to predict financial statements fraud. However, since those models based on the general abnormal accruals, those models are difficult to be applied in the banking industry. Therefore, the future research in the banking industry can focus on the association of the specific fraud incentives in the banking industry and its impact on the likelihood of financial statements fraud. The summary of the important literature is presented in Appendix 1.

3. Hypothesis development

This chapter uses the theoretical background of financial statements fraud and earnings management to develop the hypotheses. First, this chapter develops the hypothesis of the association between the financial statements fraud and prior earnings management. Next, this chapter develops the hypotheses of the association between financial statements fraud and the manipulation incentives.

3.1 Earnings management

Prior literature finds a positive and significant association between earnings management and financial statements fraud (Dechow et al., 1996; Beneish, 1999a. Dechow et al., 2011, Perols & Lougee, 2011). Since firms commonly use discretionary accrual to manage their earnings, Dechow et al. (1996) find that the discretionary accrual increase from three years before the manipulation year and decrease after the manipulation year. After examining several accrual models, Jones et al. (2008) find that discretionary accruals models can predict earnings manipulations. Following Dechow et al. (1996), Perols and Lougee (2011) predict and find that firms with greater prior earnings management are more likely to commit financial statements fraud. Using the time-series analysis, Dechow et al. (2011) suggest that before the misstatement period, managers are more likely to manage their earnings within the accounting standards.

In the banking industry, Uygur (2013) finds evidence that fraud incentives positively associated with earnings management through loan loss provision. Furthermore, Beatty and Liao (2014) find that current earnings management captured by absolute abnormal loan loss provision can predict restatements and SEC comment letters. Extending Beatty and Liao (2014) findings, and following Dechow et al. (2011) and Perols and Lougee (2011), this thesis assumes that banks opportunistic behavior captured by abnormal loan loss provision can predict manipulations captured by restatements or the SEC investigation. Furthermore, following Dechow et al. (2011) who elaborate firms' financial characteristics prior to the manipulation year, and Perols and Lougee (2011) who use prior earnings management to predict financial statements fraud, this thesis assumes that the change of prior abnormal loan loss provision positively associated with financial statements fraud. In conclusion, this thesis argues that banks are more likely to manage their earnings using loan loss provision before committing fraud.

Hypothesis 1 (H1): Banks with prior earnings management are more likely to commit financial statements fraud.

3.2 The fraud incentives

The positive accounting theory from Watts and Zimmerman (1986) suggest several hypotheses that can motivate managers' opportunistic behavior. Heally and Wahlen (1999) mention that capital market, contract, and regulatory incentives can drive managers to perform earnings management. Using the assumption that earnings manipulation shares the same objective as earnings management, Dechow et al. (1996) examine several earnings management incentives and then find that earnings management incentives can also motivate earnings manipulation.

In each incentive below, this argues that the particular incentive can motivate financial statements fraud. However, despite of the possibility that an earnings management variable can be a mediating variable between a management incentive variable and a financial statements fraud variable, following Perols and Lougee (2011), this thesis also argues that the fraud incentives can also strengthen the association (i.e. be a moderating variable) between the prior earnings management and the financial statements fraud.

3.2.1 Cost of capital

Stolowy and Breton (2004) mention that low cost of capital is one of the objectives of earnings management. Furthermore, regarding the manipulation behavior, Dechow et al. (1996) find that firms violate the regulation (e.g. accounting standards) to attract a low cost of external fund. However, Beneish (1999b) finds that the leverage of the manipulator and non-manipulator firms do not significantly different. Dechow et al. (2011) mention the inconsistency of Beneish (1999b) finding because of the observations used by Beneish (1999b). Since money becomes the commodity in the banking industry, banks have higher liabilities than the other industry. Therefore, banks' cost of capital majorly impacted by their cost of debt (i.e. interest expenses from liabilities). It can be seen in bank financial statements, interest expense is the most dominant expenses in bank income statements. Since the demand for a low-cost financing can motivate firms to manipulate their earnings (Dechow et al., 1996), this thesis argues that bank cost of capital has a positive and significant association with financial statements fraud.

Hypothesis 2.1.a (H21a) Banks with a higher cost of capital are more likely to commit financial statements fraud

Next, because firms can manage their earnings to attract a low cost of capital (Stolowy and Bretton, 2004), this thesis also argues that during a high cost of capital circumstance banks with prior earnings management are more likely to commit financial statements fraud the more the banks have a higher cost of capital.

Hypothesis 2.1.b (H21b) Banks with prior earnings management are more likely to commit financial statements fraud the more they have higher interest expenses.

3.2.2 Minimum capital requirement

The minimum capital requirement is one of the examples of the industry-specific regulation that can motivate earnings management (Heally & Wahlen, 1999). Including in the CAMELS bank ratings, the C or the capital minimum requirement is one of the bank soundness indicators (Beatty & Liao, 2014). There are several consequences of a lower amount of capital requirements such as slower assets growth (Beatty and Liao, 2011) and regulatory scrutiny. Since the minimum capital requirement ratio based on the accounting numbers and earnings is a part of the capital (i.e. tier-1 capital), banks overstate their earnings to increase their capital. To increase their earnings, banks understates their loan loss provision (Ahmed et al., 1999; Ronen & Yaari, 2008; Beatty & Liao, 2014). Based on those findings, this thesis has two arguments. First, this thesis argues that during low CAR circumstances banks are more likely committing financial statements fraud. Second, this thesis argues that during low CAR circumstances banks with prior earnings management are more likely to commit financial statements fraud.

Hypothesis 2.2.a (H22a) Banks with a lower amount of the capital requirement are more likely to commit financial statements fraud.

Hypothesis 2.2.b (H22b) Banks with prior earnings management are more likely to commit financial statements fraud the more they have a lower amount of the capital requirement.

3.2.3 Bank distress

DeAngelo, DeAngelo, and Skinner (1994) mention that managers are more likely to manage their accounting numbers due to financial distress than to overstate their earnings. Beneish (1997) suggest that firms with extreme financial condition (e.g. financial distress) are more likely to violate the accounting standards. Furthermore, Rosner (2003) finds that firms tend to manipulate their reports during a pre-bankruptcy period. In the banking industry, Chiaramonte et al. (2016) find that financial distress indicators can predict bank failure. Therefore, since banks have more incentives to manipulate their earnings during distress period, this thesis argues that banks with higher distress circumstances are more likely to commit financial statements fraud.

Hypothesis 2.3.a (H23a) Banks with higher financially distress are more likely to commit financial statements fraud.

Since DeAngelo et al. (1994) mention that managers engaged with earnings manipulation in the financial distress condition, this also argues that during distress circumstances, banks with prior earnings management are more likely to commit financial statements fraud.

Hypothesis 2.3.b (H23b) Banks with prior earnings management are more likely to commit financial statements fraud the more they faced higher distress.

3.2.4 Liquidity

Beneish (1997:1999a) show that firms with lower liquidity are more likely to manipulate their earnings. Dechow et al. (2011) find that lack of liquidity increase the probability of accounting misstatements. In the banking industry, liquidity issues create severe damages, lower liquidity issues harm bank reputation and attract more regulators attention. The recent study from Chiaramonte and Casu (2016) show that less liquid banks are more likely to fail than more liquid banks. However, with also considering other perspectives that banks with higher liquidity also have a higher opportunity cost (i.e. inefficiency), and otherwise that banks with higher liquidity have a higher reputation, and therefore tend to have less expensive fund. First, this thesis argues that banks with lower liquidity are more likely committing financial statements fraud.

Hypothesis 2.4.a (H22a) Banks with lower liquidity are more likely to commit financial statements fraud.

Next, since firms are more likely engaged in earnings management before committing financial statements fraud (Perols and Lougee, 2011), this thesis argues that during low liquidity circumstance banks with prior earnings management are more likely to commit financial statements fraud.

Hypothesis 2.4.b (H24b) Banks with prior earnings management are more likely to commit financial statements fraud the more they have lower liquidity.

4. Research design

This chapter explains the statistical method to test the hypotheses and the sample selection process. First, this chapter discusses the main model to test the hypotheses and then the variables used in the model. The next section explains the reasons for the observations used, and then the steps to get the final sample.

4.1 Model to test the hypotheses

Following Dechow et al. (2011), Perols and Lougee (2011), and Beatty and Liao (2014), I use logit regression to find the association between financial statements fraud with prior earnings management and the fraud incentives. In order to test the first hypothesis and the hypotheses 2.a (i.e. the fraud incentives), this thesis uses equation 8. Regarding the hypotheses 2.b, the fraud incentives as the moderating variables, this thesis uses equation 9. The dependent variable, financial statements fraud (FSF) is a dummy variable, equal to 1 if a bank commits financial statements fraud and 0 otherwise. The independent variables: Prior earnings management is abnormal loan loss provision and delayed loan loss provision; Cost of capital (CC) is interest expenses divided by total liabilities; Capital adequacy ratio (CAR) is tier-1 capital divided by risk-weighted assets; Bank distress is bank Z-score; and Liquidity is cash and equivalents divided by total assets. The control variables are the audit firm (BIG4), total assets (SIZE), profitability (ROA), and internal control material weakness (ICMW). In the next section, I explain the variables in more detail.

Equation 7

$$FSF_{j,t} = \beta_0 + \beta_1\, Prior\ earnings\ management + \beta_2\, Cost\ of\ capital + \beta_3\, CAR + \beta_4\, Bank\ distress + \beta_5\, Liquidity + \beta_n\, control\ variables + \varepsilon$$

Equation 8

$$\begin{aligned} FSF_{j,t} = {} & \beta_0 + \beta_1\, Prior\ earnings\ management \\ & + \beta_2\, Cost\ of\ capital + \beta_3\, CAR + \beta_4\, Bank\ distress + \beta_5\, Liquidity \\ & + \beta_6\, Prior\ earnings\ management * Cost\ of\ capital \\ & + \beta_7\, Prior\ earnings\ management * CAR \\ & + \beta_8\, Prior\ earnings\ management * Bank\ distress \\ & + \beta_9\, Prior\ earnings\ management * Liquidity + \beta_n\, control\ variables + \varepsilon \end{aligned}$$

The logit regression in equation 8 is used to test the first and the second hypotheses (i.e. hypotheses 2.a). Regarding the prior earnings management and the manipulation incentives hypotheses, H1 and H21a predict that β_1 and β_2 respectively significant and positive. Next, H22a, H23a, and H24a respectively predict that β_3, β_4, and β_5 significant and negative.

Equation 9 is used to test the hypotheses 2.b (i.e. fraud incentives as moderating variables). Regarding the moderating variables hypotheses, H21b, H22b, H23b, and H24b respectively predict β_6, β_7, β_8, and β_9 significant and positive. The predictive validity framework (Libby boxes) is presented in Appendix 2.

4.2 Variables explanation

4.2.1 Financial statements fraud

In the theoretical background, section 2.1.1.4, this thesis discusses why researchers commonly use SEC AAERs and accounting restatements as the proxy of financial statements fraud or accounting misstatements. Researchers use both SEC AAERs and accounting restatements as a direct proxy of financial statements fraud. SEC AAERs commonly contain firms with a strong indication of financial statements fraud, or high type 1-error. However, because SEC commonly investigates high profile violation cases, SEC AAERs potentially excludes relatively low impact financial statements fraud or low type-2 error. In the banking industry, Beatty and Liao (2014) also use financial restatements and SEC comment letters related to loan loss provision to capture provision manipulation.

Following Perols and Lougee (2011) and Perols et al. (2016), and therefore inconsistent with Dechow et al. (2011), this thesis use term “financial statements fraud” and not accounting misstatements or misreporting. In the same logic, this thesis uses Dechow et al. (2011) explanation that managers commonly deny their fraud intention. In addition, it is also difficult to find solid reasons that managers do not realize the significant error in the financial statements. Therefore, this thesis excludes banks with error and clerical restatements from the fraud observations. In addition to auditor restatements or SEC restatements, following Costello et al. (2016) who use regulators reports (i.e. catch-up reports), this thesis also uses regulators restatements as the proxy of financial statements fraud.

4.2.2 Earnings management

Abnormal loan loss provision captured by the residual of loan loss provision accrual model commonly used by researchers to measure banks' earnings management behavior (e.g. Ma, 1988; Beatty et al., 2002; Bushman & Williams, 2012; Uygur, 2013; Beatty and Liao, 2014). However, currently, researchers examine the use of delayed loan loss recognition model to measure bank opacity (Nichols et al., 1999; Beatty & Liao, 2011; Bushman & Williams, 2015). Bushman and Williams (2015) suggest that delayed loan loss provision can capture bank opportunistic behavior. Therefore, in addition to abnormal loan loss provision, this thesis also uses delayed loan loss recognition as the proxy of earnings management.

4.2.2.1 Prior earnings management through abnormal loan loss provision

First, since this thesis follows Dechow et al. (2011) who examine the period prior to the misstatements period, and Perols and Lougee (2011) who use the last three years' abnormal accrual to predict financial statements fraud, this thesis measures bank prior earnings management as the increase of absolute abnormal loan loss provision in the last three years. Since bank loan quality is the main driver of both interest revenue and loan loss provision, the major impairment in loan quality can significantly reduce bank financial condition. Therefore, this thesis assumes that in the condition of loan quality impairment, banks can only maintain their loan quality for the last two years (i.e. eight quarters period). In that period, banks manage their earnings through abnormal loan loss provision. Furthermore, as banks financial condition decrease, banks increase their earnings management before the banks committing fraud.

Equation 10

$$Prior\ earnings\ management_{j,t} = \epsilon_{j,t} - \epsilon_{j,t-8}$$

Bank j prior earnings management in quarter t is equal to the average value of absolute abnormal loan loss provision in quarter t minus the average value of absolute abnormal loan loss provision eight quarters prior to the t quarter ($\varepsilon_{j,t-8}$).

Second, following Beatty and Liao (2014), this thesis uses the average values of absolute abnormal loan loss provision. The average value is a bank mean abnormal loan loss provision in one year. Despite using the real value of abnormal loan loss provision as it captured by the residual of loan loss provision model, the absolute value can capture both the greater negative and the greater positive abnormal loan loss provision. Ma and Song (2016) mention that the higher the absolute value of abnormal loan loss provision, the higher the magnitude of earnings management.

Third, the abnormal loan loss provision in this thesis is the residual value of loan loss provision accrual model from Beatty and Liao (2014). I use Beatty and Liao (2014) loan loss provision accrual models (equation 1) because they develop their models based on the evaluation of the prior loan loss provision accrual models. To compare the abnormal loan loss provision from Beatty and Liao (2014) models, this thesis also uses abnormal loan loss provision value from Bushman and Williams (2012) model (equation 3). However, following Beatty and Liao (2014), in the main analysis, this thesis excludes the earnings before loan loss provision and tier-1 capital requirement from Bushman and Williams (2012) model. Because this thesis measures opportunistic earnings management, this thesis excludes the earnings before loan loss provision since it captures earnings smoothing behavior. Next, this thesis also excludes the minimum capital requirement variable since the variable is one of the fraud variables which also measure in this thesis. This thesis includes the both variables in the additional test, section 5.4.1.1.2.

4.2.2.2 Delayed loan loss recognition

Timely loan loss recognition or small delay on loan loss recognition means banks have anticipated the potential loan quality decrease in the future (Bushman & Williams, 2015). Therefore, if banks know that the loan quality will decrease in the next reporting period, but the banks do not want to increase their loan loss provision because it can decrease their current earnings, the banks then manage the current period earnings by delaying the recognition of loan loss. Furthermore, Bushman and Williams (2015) suggest that opportunistic or more aggressive banks tend to defer the recognition of loan loss and this positively associated with bank opacity.

Following Bushman and Williams (2015), this thesis calculates delayed loan loss recognition from the difference between adjusted R^2 from equation 4 model (1) and model (2). The incremental R^2 of each observation than compares with the medium of values of each quarter incremental R^2. Observations with lower incremental R^2 than the medium R^2 are then classified as banks with higher delayed on loan loss recognition.

Equation 9

$$\textit{Incremental } R^2 = \textit{equation 4, model (2) adjusted } R^2 - \textit{equation 4, model (1) adjusted } R^2$$

4.2.3 The fraud incentives

4.2.3.1 Cost of capital

Bank cost of capital majorly impacted by bank cost of debt. Interest expense is the greatest part of bank total expenses. Since higher interest expense means higher cost of debt, I use the proxy from Betz et al. (2014) to measure bank cost of capital.

Equation 10

$$Cost\ of\ capital_{j,t} = \frac{Interest\ expenses_{j,t}}{Total\ liabilities_{j,t}}$$

Based on the hypothesis 2.1.a, the prediction is the higher the cost of capital, the higher the financial statements fraud probability.

4.2.3.2 Capital requirement

The minimum capital requirement ratio is from the Basel framework. Tier-1 capital divided by risk-weighted assets is available in Compustat bank database.

Equation 11

$$CAR_{j,t} = \frac{Tier - 1\ capital_{j,t}}{Risk - weighted\ assets_{j,t}}$$

Based on the hypothesis 2.2.a, the prediction is the lower the CAR, the higher the financial statements fraud probability.

4.2.3.3 Bank distress

Chiaramonte et al. (2016) examine the US bank data from the year 2004-2012 and find that several bank Z-score can predict bank failure. Following Chiaramonte et al. (2016), this thesis uses the Z-score model from DeLisle, Maecheler, & Srobona (2007) to measure bank distance to failure.

Equation 12

$$Bank\ distress_{j,t} = \frac{3\ years\ moving\ average \frac{Equity_{j,t}}{Total\ assets_{j,t}} + 3\ years\ moving\ average\ ROA_{j,t}}{3\ years\ \sigma\ ROA_{j,t}}$$

Based on hypothesis 2.3.a, the prediction is the lower the bank distress, the higher the financial statements fraud probability.

In the additional test section, this thesis also uses the other three Z-score model examined by Chiaramonte et al. (2016).

4.2.3.4 Liquidity

Kim and Sohn (2017) who examine the US banks liquidity used cash and securities divided by total assets to measure bank liquidity. In line with them, Aydemir and Guloglu (2017) measure liquidity risk as liquid assets divided by total assets. Following the prior literature, this thesis uses cash and equivalents divided by total assets as the liquidity variable.

Equation 13

$$Liquidity_{j,t} = \frac{Cash\ and\ cash\ equivalent_{j,t}}{Total\ assets_{j,t}}$$

Based on hypothesis 2.4.a, the prediction is the lower the liquidity, the higher the financial statements fraud probability.

In the additional test section, this thesis also uses the net stable funding ratio (NSFR) as the proxy of banks liquidity.

4.2.4 Control variables

4.2.4.1 The audit firm

Fanning and Cogger (1998) find that the auditor quality negatively associated with fraud. In the banking industry, Ma and Song (2016) find that the effect of earnings management on bank systemic risk decrease if the bank is audited by one of the Big 4 audit firms. To control the audit quality, I use the audit firm as one of the control variables. The BIG4 is a dummy variable equal to 1 if the bank auditor is from the Big 4 audit firms and equal to 0 if otherwise. The prediction is banks with the Big 4 audit firms are less likely to commit financial statements fraud.

4.2.4.2 Bank size

In the banking industry, large banks are more associated with earnings management through loan loss provision (Uygur, 2013). To control the bank size, I use the natural log of bank total assets (*SIZE*) as one of the control variables. The prediction is banks with greater total assets are more likely to commit financial statements fraud.

4.2.4.3 Return on assets

Dechow et al. (2011) find that firms with decreasing return on assets are more likely engaged with accounting misstatements. Perols and Lougee (2011) find that firms with lower profitability are more associated with financial statements fraud. In the banking industry, Uygur (2013) finds that bank earnings management using loan loss provision negatively associated with the return on assets (ROA). Therefore, I use the return on assets (*ROA*) as one of the control variables. The prediction is banks with lower ROA are more likely to commit financial statements fraud.

4.2.4.4 Internal control

Doyle et al. (2007) find that firms with internal control material weakness (ICMW) are more likely to have higher earnings management. Furthermore, they find that internal control weakness has a negative association with the governance structure. In line with Doyle et al. (2007), Dechow et al. (2010) suggest that internal control deficiency can also indicate accounting misstatements. Since internal control can also capture firms' governance, this thesis use internal control opinion from bank auditor to control the governance structure. The ICMW is a dummy variable equal to 1 if there is ICMW and equal to 0 if otherwise. The prediction is banks with ICMW are more likely to commit financial statements fraud.

4.3 Data & sample selection

This thesis examines the United States banks data from the year 2003-2015 or after the implementation of Sarbanes-Oxley Act. Since the SOX act encourages companies to increase their governance and to mitigate fraud, there is a difference between managers' opportunistic behavior before and after the SOX implementation. This thesis uses Bank Holding Company (BHC) as my sample because most of the public banks are holding companies. Following Dechow et al. (2011) and therefore inconsistent with Beneish (1999a) and Perols and Lougee (2011) who match the fraud firms with the selected non-fraud firms, this thesis randomly selects the fraud and the non-fraud observations.

The majority of the data are from Wharton Research Data Services (WRDS) system which is subscribed by the University Library. The financial statements fraud variable is from Bank Regulatory and Audit Analytics. The financial data to calculate most of the independent variables are from Compustat Bank. Both control variables which are the audit firm (*BIG4*) and the internal control material weakness (*ICMW*) are from Compustat North America. Finally, the macroeconomic variables are from Federal Reserve St. Louis and Bureau of Labor Statistics websites.

Before merging the datasets, this thesis calculates the earnings management and the other financial variables from Compustat Bank dataset. Because this thesis uses future and lagged variables, calculating the financial variables in the first place can avoid missing values in my observations. There are 35,759 bank-quarter observations after excluding observations outside the period. In the merging process, this thesis uses the Compustat Bank as the main table. First, this thesis merges the main table with the Unemployment, the GDP, and the Case-Shiller tables. Next, using CUSIP and quarter as the key variables, this thesis merges the main table with the Bank Regulatory table. Because CUSIP consistently available in Bank Regulatory dataset since 2006, I use the bank specific identifier in the Bank Regulatory to get the CUSIP identifier from the year 2003-2005. Next, this thesis merges the main table with Compustat North America table. Because Audit Analytics table only contains firms with restatements, this thesis adds restatements from the Audit Analytics table as a new variable in the main table. Regarding the time variable in the Audit Analytics, Non-reliance Restatements table, this thesis uses auditor restatements filing date as the quarter period. Since this thesis focusses on the fraud instead of errors, from Audit Analytics table, this thesis excludes banks with error and clerical restatements. Finally, the merging process results in 11,365 bank-quarter observations. After drop missing and duplicate variables, and then winsorizing the loan loss provision variable[3], the final sample contains 9,715 bank-quarter observations that represent 420 banks. The overall observations consist of 1,073 fraud observations and 8,642 non-fraud observations. The fraud observations are the mix of 968 regulator restatements, 119 auditor accounting restatements, 1 fraud restatements, and 6 SEC investigation restatements.

Table 1
Sample Selection Process

Compustat Bank			35,759
Merging datasets			
- Unemployement, GDP, and Case-Shiller tables	-368		
- Compustat North America	-13,791		
- Bank Regulatory-BHC*, & Audit Analytics-NRR* tables	-10,235		
Less total merging		-24,394	
Observations after merging			11,365
Missing and duplicate variables		-1,650	
Final observations			9,715

BHC: Bank Holding Companies, NRR: Non-Reliance Restatements

[3] Same with Beatty and Liao (2014) sample, the loan loss provision values in this thesis are also skewed.

5. Results

This chapter discusses the statistical results to answer the hypotheses. First, this chapter shows the descriptive and time-series analyses. Next, this chapter shows the regression results. Finally, this chapter shows the additional tests and the prediction models.

5.1 Descriptive analysis

Table 2 classifies the observations into two groups which are the fraud and the non-fraud groups. The table shows the mean value of the variables and compares it between the two groups. The last two columns are the prediction based on the theoretical background and the p-value of the mean differences.

Table 2 Descriptive statistics

	Full sample mean	Fraud (F) mean	Non-Fraud (NF) mean	Difference (3) - (2)	Prediction	Difference P-value
NPA	0.0220	0.0212	0.0221	0.0009	F>NF	0.2749
LLP	0.0018	0.0027	0.0017	-0.0011	F>NF	0.0000***
Abnormal LLP	0.0014	0.0017	0.0014	-0.0003	F>NF	0.0000***
Prior earnings management (Hypothesis 1)						
Abnormal LLP (Δ3y)	0.0000	0.0006	-0.0001	-0.0007	F>NF	0.0000***
Fraud incentives (Hypotheses 2a)						
CC	0.0043	0.0061	0.0041	-0.0020	F>NF	0.0000***
CAR	12.2611	11.1215	12.4026	1.2811	F<NF	0.0000***
B_DIS	231.6180	203.9502	235.0753	31.1251	F<NF	0.0000***
LIQ	0.0415	0.0413	0.0415	0.0001	F<NF	0.9157
Control variables						
ROA	0.0018	0.0010	0.0020	0.0010	F<NF	0.0000***
SIZE	8.2199	8.9747	8.1262	-0.8485	F>NF	0.0000***
Growth	0.0731	0.0657	0.0736	0.0079	F>NF	0.4214
N	9715	1073	8642	9715		

*** $p < 0.01$

Definition of the variables:

NPA	Non-performing assets divided by total loan
LLP	Total loan loss provision divided by lagged total loan
Abnormal LLP	The average abnormal loan loss provision is from the residual value of Beatty and Liao (2014) loan loss provision accrual model (a)
Abnormal LLP(Δ3y)	The change of model (a) abnormal loan loss provision for the last three years.
CC	Cost of capital: interest expenses divided by total loans
CAR	Tier 1 capital divided by weighted average asset risk (%)
B_DIS	Bank distress: DeLisle et al. (2007) Z-score.
LIQ	Liquidity: cash and equivalents divided by total assets
ROA	Return on assets: income divided by average total assets
SIZE	Natural log of total assets
Growth	One-year total liabilities growth

First, regarding the non-performing assets (NPA), there is no a significant difference between the fraud and the non-fraud groups. However, the fraud group loan loss provision (LLP) significantly higher than the non-fraud group. The real earnings management practice can be the reason why fraud group has lower non-performing assets. Furthermore, lower non-performing assets with higher loan loss provision can also be an indicator of poor credit risk management. Next, fraud group has higher abnormal loan loss provision than the non-fraud group. Regarding the first hypothesis, compare to the non-fraud group, the fraud group has higher abnormal loan loss provision increase in the last three years[4].

Next, regarding the second hypotheses, in line with Dechow et al. (1996), the fraud group has a relatively higher cost of debt (CC) than the non-fraud group. Consistent with Heally and Wahlen (1999) and Beatty and Liao (2014), observations with a lower minimum capital requirement (CAR) has higher incentive to commit financial statements fraud. Next, banks with lower Z-score or higher failure risk are more likely to commit fraud. Inconsistent with the liquidity assumption in the second hypothesis, there is no significant difference between the liquidity of fraud and non-fraud groups. Finally, regarding the control variables, consistent with Beneish (1999a) and Perols and Lougee (2011), the fraud group has a lower return on assets and higher size than the non-fraud group. However, inconsistent with Beneish (1999a), Perols and Lougee (2011), and Uygur (2013), there is no significant difference between the fraud and the non-fraud groups' growth of liabilities.

5.1.1 Fraud distribution and prior earnings management

Table 3 shows fraud and non-fraud observations distribution per year. In total, there are 1,073 fraud observations that represent 11.04% of the total sample. More than half of the fraud observations are from the years around the financial crisis (2006-2009). Regarding the fraud frequency, the table shows that in the financial crisis years the fraud frequency is higher than the other periods. Regarding the prior earnings management, table 3 shows that around the financial crisis years, the prior earnings management are significantly higher than the other years. In general, the fraud group prior earnings management still higher than the non-fraud group. Since the banking industry is one of the industries that impacted the most by the financial crisis, the table shows that financial crisis gives more incentives to the manager to perform both financial statements fraud and earnings management.

[4] The loan loss provision accrual model to generate abnormal loan loss provision in table 3 is discussed in section 5.3.1.

Table 3 Fraud distribution and prior earnings management

Year	Fraud frequency				Prior earnings management				
	Full Sample	Fraud	Non-fraud	%Fraud	Full Sample	Fraud	Non-fraud	Difference	P-value
2003	502	18	484	3.59%	0.00008	0.00036	0.00007	0.00029	0.3650
2004	427	13	414	3.04%	-0.00050	-0.00029	-0.00051	0.00022	0.5262
2005	702	20	682	2.85%	-0.00049	-0.00015	-0.00050	0.00035	0.2891
2006	946	106	840	11.21%	-0.00055	-0.00056	-0.00054	-0.00002	0.8822
2007	921	219	702	23.78%	0.00007	0.00021	0.00003	0.00017	0.0705
2008	882	351	531	39.80%	0.00129	0.00121	0.00134	-0.00013	0.2112
2009	806	143	663	17.74%	0.00210	0.00251	0.00201	0.00049	0.0041
2010	751	54	697	7.19%	0.00062	-0.00043	0.00071	-0.00114	0.0001
2011	732	17	715	2.32%	-0.00110	-0.00108	-0.00110	0.00002	0.9733
2012	734	30	704	4.09%	-0.00073	-0.00099	-0.00072	-0.00027	0.4954
2013	775	10	765	1.29%	-0.00023	-0.00050	-0.00023	-0.00027	0.5364
2014	779	38	741	4.88%	-0.00058	-0.00050	-0.00058	0.00008	0.6406
2015	758	54	704	7.12%	-0.00047	-0.00057	-0.00047	-0.00011	0.3276
Total	9,715	1,073	8,642	11.04%	-0.00004	-0.00006	-0.00004	-0.00002	0.0000

prior earnings management equal to the increase of the average value of Beaty and Liao (2014) model (a) absolute abnormal loan loss provision in the last three year

5.2 Time-series analysis

Following Dechow et al. (2011), this section compares the changes of several important variables between the fraud and the non-fraud groups. The figures show the variables changes eight quarters before the fraud period to four quarters after the fraud quarters. First, this section discusses the non-performing assets and the abnormal loan loss provision (hypothesis 1), and then the fraud incentives variables (hypotheses 2).

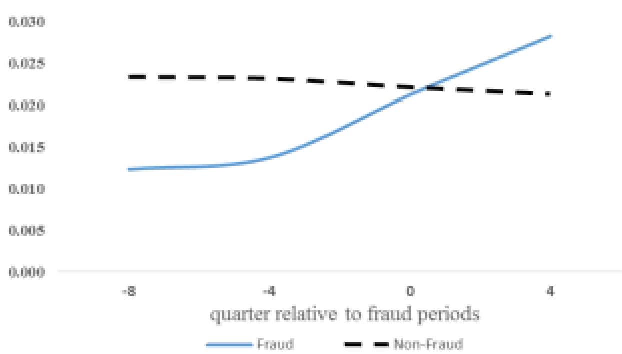
5.2.1 Non-performing assets and loan loss provision

Figure 1.a shows that two years before the fraud period, the non-performing assets of the fraud group is lower than the non-fraud group. However, after the fraud period, the fraud group non-performing assets increase higher than the non-fraud group. The possible reason of this result is fraud banks are more likely to manage their reports with postponed the loan quality impairment i.e. earnings management through loan restructuring. However, after the auditing process by the regulators or the public accountants, the banks are required to restate their reports, including to reveal their true non-performing assets. In line with the non-performing assets increase, figure 1.b shows that the total loan loss provision of the fraud group increase gradually and then becomes higher than the non-fraud group. It also means that the fraud group has known the risk of loan loss impairment in the future.

The absolute value of abnormal loan loss provision in figures 1.c and 1.d are respectively from Beatty and Liao (2014) model (a) and Bushman and Williams (2012)[5]. Beatty and Liao (2014) and Ma and Song (2016) mention that the higher the value of absolute abnormal loan loss provision, the higher the magnitude of the opportunistic earnings management. Therefore, figures 1.c and 1.d support the first hypothesis.

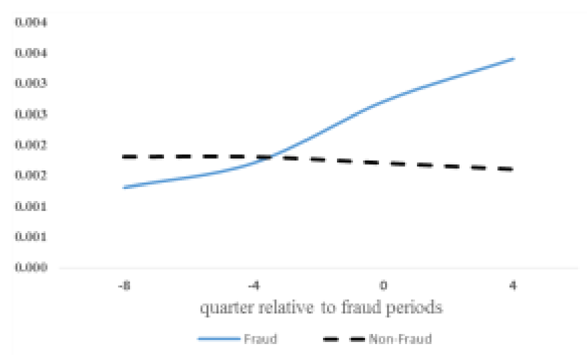
Figure 1
Non-performing assets and loan loss provision changes

Figure 1.a Non-performing assets (npa)



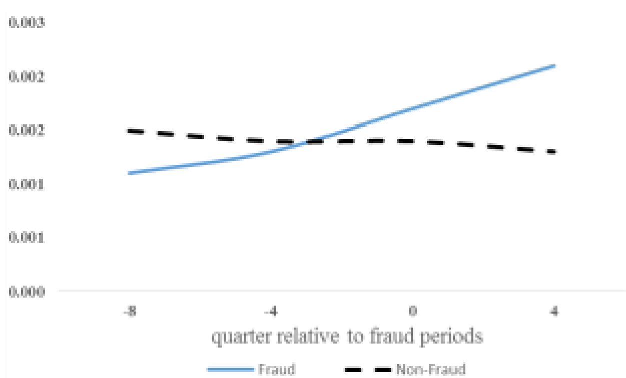
npa: non-performing assets divided by total loan

Figure 1.b Total loan loss provision (LLP)



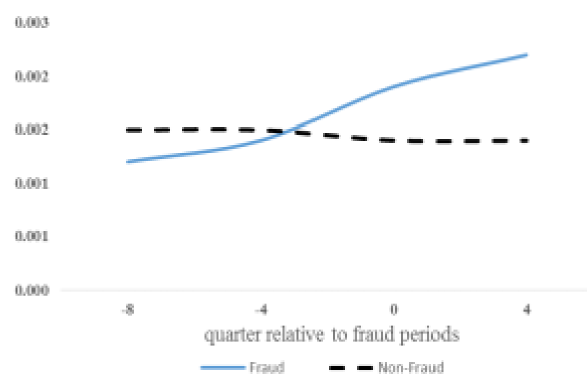
LLP: loan loss provision divided by lagged total loans

Figure 1.c
Model (a) abnormal loan loss provision



Model (a): Beatty and Liao (2014) loan loss provision model (a)

Figure 1.d
Model (e) abnormal loan loss provision



Model (e): Bushman and Williams (2012) loan loss provision model

[5] The loan loss provision models to generate abnormal loan loss provisions in figure 1.c and figure 1.d are discussed in section 5.3.1.

5.2.2 Fraud incentives variables

Figures 2.a and 2.b show that around the fraud period, fraud group cost of capital and minimum capital requirement respectively higher and lower than the non-fraud group. These two figures support the hypotheses 2.a and 2.b. Next, figure 2.c shows that before the fraud period, the fraud group has lower financial distress than the non-fraud group. However, the conditions before the fraud period potentially do not reflect banks' true condition due to the real earnings management practices. This reason can explain why after the fraud period the fraud group financial distress increase and then higher than the non-fraud group. Figure 2.d shows that only before the fraud period the fraud group liquidity is higher than the non-fraud group liquidity. In general, the figures show that that examining the period around the fraud period is important to know the fraud incentives behavior.

Figure 2
The fraud incentives changes

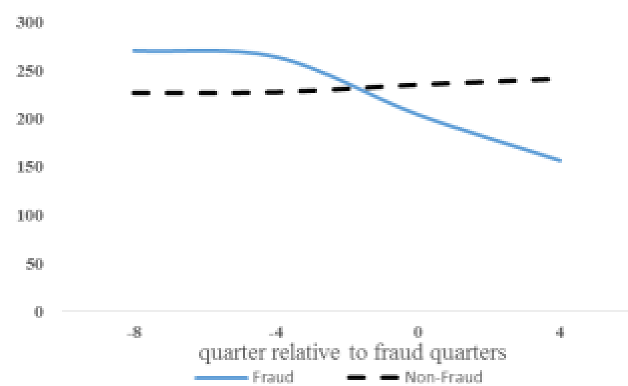
Figure 2.a. Cost of capital

Cost of capital: interest expenses divided by total loans.

Figure 2.b. CAR

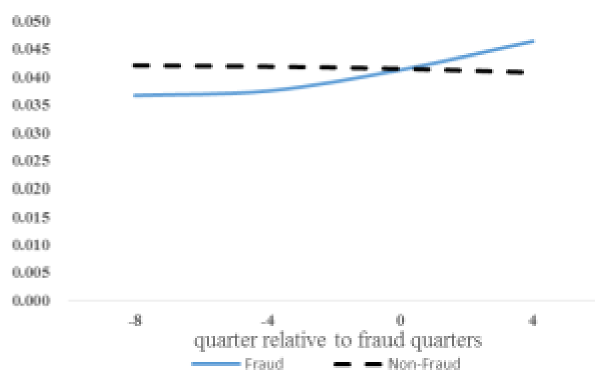
CAR: Tier-capital divided by weighted average asset risk (%).

Figure 2.c. Bank distress



Bank distress: DeLisle et al. (2007) Z-score.

Figure 2.d. Liquidity



Liquidity: Cash and equivalents divided by total assets.

Multivariate analysis

5.3.1 Loan loss provision accrual model

5.3.1.1 Ordinary least square

Table 4 shows the results of ordinary least square (OLS) regression of loan loss provision accrual model. Model (a) – model (d) are the main models from Beatty and Liao (2014). I add model (e) from Bushman and Williams (2012) to compare the results of the main models. I choose Bushman and Williams (2012) model for several reasons: First, they develop their model using relatively new datasets compare to the other LLP models. Next, they use their model to examine the likelihood of earnings management after the future implementation of expected credit losses. All models in table 4 use future, current, and prior non-performing assets to predict loan loss provision. Together, they also use change in GDP as the macroeconomic variable in their models. However, different to Bushman and Williams (2012) model, Beatty and Liao (2014) models use the change of total loans, loan loss allowance and charge-off to predict bank loan loss provision.

Regarding the results, consistent with Beatty and Liao (2014), except for the bank size ($SIZE_{t-1}$) and the change of total loans (ΔLoan) in the model (c) and the model (d), all variables have a significant association with loan loss provision. Table 4 shows as the number of the independent variables increase from model (e) to model (d), the adjusted R-square of the model increase. More specific, the models show that the next year (ΔNPA_{t+1}), the current (ΔNPA_{t}), and the prior years' (ΔNPA_{t-1} & ΔNPA_{t-2}) change of non-performing assets are important elements of loan loss provision recognition. The association between next year non-performing loan captures managers' anticipation of future loan quality impairment, and the association between the last two years non-performing assets represents the fact that banks use historical analysis of loan impairment to predict current loan loss provision.

Table 4 OLS regression of loan loss provision

Variables	Prediction	Model (a) Coefficients (t-statistics)	Model (b) Coefficients (t-statistics)	Model (c) Coefficients (t-statistics)	Model (d) Coefficients (t-statistics)	Model (e) Coefficients (t-statistics)
Dependent variable: LLP						
ΔNPA_{t+1}	+	0.022^{***}	0.030^{***}	0.023^{***}	0.022^{***}	0.013^{***}
		(6.44)	(8.85)	(6.86)	(6.66)	(2.67)
ΔNPA	+	0.051^{***}	0.059^{***}	0.044^{***}	0.043^{***}	0.044^{***}
		(5.46)	(5.67)	(4.21)	(4.08)	(5.32)
ΔNPA_{t-1}	+	0.046^{***}	0.052^{***}	0.028^{***}	0.027^{***}	0.044^{***}
		(10.02)	(10.77)	(8.40)	(8.29)	(9.14)
ΔNPA_{t-2}	+	0.047^{***}	0.052^{***}	0.024^{***}	0.022^{***}	0.048^{***}
		(10.26)	(11.07)	(6.56)	(6.39)	(9.12)
$SIZE_{t-1}$	+	0.000^{***}	0.000^{***}	0.000	0.000	0.000^{***}
		(4.05)	(3.98)	(0.73)	(0.93)	(3.87)
ΔLoan	+	-0.012^{***}	-0.006^{***}	0.001	0.000	
		(-8.47)	(-5.62)	(1.03)	(0.46)	
ΔGDP	-	-0.021^{***}	-0.021^{***}	-0.009^{***}	-0.009^{***}	-0.038^{***}
		(-11.03)	(-11.35)	(-7.80)	(-7.63)	(-15.78)
CSRET	-	-0.004^{***}	-0.002^{***}	-0.000^{***}	-0.001^{***}	
		(-18.28)	(-7.66)	(-3.26)	(-4.53)	
ΔUNEMP	+	0.002^{***}	0.002^{***}	0.002^{***}	0.002^{***}	
		(5.07)	(5.97)	(6.66)	(6.56)	
ALW_{t-1}	+		0.119^{***}		-0.017^{**}	
			(13.09)		(-2.52)	
CO	+			0.837^{***}	0.866^{***}	
				(42.60)	(40.23)	
constant	?	0.008^{***}	0.003^{***}	0.001^{***}	0.002^{***}	0.002^{***}
		(15.69)	(5.10)	(6.35)	(7.36)	(4.78)
R-sqr		0.315	0.405	0.740	0.741	0.232
N		9715	9715	9715	9715	9715

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Standard errors are clustered at the bank level.

Model (a). $LLP_{j,t} = \alpha_0 + \alpha_1 \Delta NPA_{t+1} + \alpha_2 \Delta NPA_t + \alpha_3 \Delta NPA_{t-1} + \alpha_4 \Delta NPA_{t-2} + \alpha_5 \Delta SIZE_{t-1} + \alpha_6 \Delta Loan_t + \alpha_7 \Delta GDP_t + \alpha_8 CSRET_t + \alpha_8 \Delta UNEMP_t + \varepsilon_{j,t}$

Model (b): $LLP_{j,t} = \alpha_0 + \alpha_1 \Delta NPA_{t+1} + \alpha_2 \Delta NPA_t + \alpha_3 \Delta NPA_{t-1} + \alpha_4 \Delta NPA_{t-2} + \alpha_5 \Delta SIZE_{t-1} + \alpha_6 \Delta Loan_t + \alpha_7 \Delta GDP_t + \alpha_8 CSRET_t + \alpha_8 \Delta UNEMP_t + \alpha_{10} \Delta ALW_{t-1} + \varepsilon_{j,t}$

Model (c): $LLP_{j,t} = \alpha_0 + \alpha_1 \Delta NPA_{t+1} + \alpha_2 \Delta NPA_t + \alpha_3 \Delta NPA_{t-1} + \alpha_4 \Delta NPA_{t-2} + \alpha_5 \Delta SIZE_{t-1} + \alpha_6 \Delta Loan_t + \alpha_7 \Delta GDP_t + \alpha_8 CSRET_t + \alpha_8 \Delta UNEMP_t + \alpha_{10} \Delta CO_t + \varepsilon_{j,t}$

Model (d): $LLP_{j,t} = \alpha_0 + \alpha_1 \Delta NPA_{t+1} + \alpha_2 \Delta NPA_t + \alpha_3 \Delta NPA_{t-1} + \alpha_4 \Delta NPA_{t-2} + \alpha_5 \Delta SIZE_{t-1} + \alpha_6 \Delta Loan_t + \alpha_7 \Delta GDP_t + \alpha_8 CSRET_t + \alpha_8 \Delta UNEMP_t + \alpha_{10} \Delta ALW_{t-1} + \alpha_{11} CO_t + \varepsilon_{j,t}$

Model (e): $LLP_{j,t} = \alpha_0 + \alpha_1 \Delta NPA_{t+1} + \alpha_2 \Delta NPA_t + \alpha_3 \Delta NPA_{t-1} + \alpha_4 \Delta NPA_{t-2} + \alpha_5 \Delta SIZE_{t-1} + \alpha_6 \Delta GDP_t + \varepsilon_{j,t}$

Definition of the variables:

LLP	the loan loss provision divided by lagged total loans.
ΔNPA	the change of non-performing assets divided by lagged total loans.
SIZE	the natural log of total assets.
ΔLoan	the change of total loan divided by lagged total loans.
ΔGDP	the change of Gross Domestic Product over the quarter.
CSRET	the return of the Case-Shiller Real Estate Index over the quarter.
ΔUNEMP	the change of unemployment rates over the quarter.
ALW	the loan loss allowance divided by total loan.
CO	the net charge-off divided by lagged total loan.

Next, table 4 shows that the control variable, the lagged total assets (SIZE) has a positive and significant association with the loan loss provision in the model (a), the model (b), and the model (e). This means that the bank size variable can capture banks risk as it shows by the loan loss provision. However, similar with Beatty and Liao (2014) results, the variable coefficient slightly above zero. Next, in line with Beatty and Liao (2014), the change of total loans (ΔLoan) negatively associated with the loan loss provision in the model (a) and the model (b).

The macroeconomic variables which are the change in gross domestic products, the return of Case-Shiller index and the change of unemployment rates have a significant association with the loan loss provision recognition. Since sound macroeconomic condition positively associated with business growth, it negatively associated with loan quality impairment. Consistent with Beatty and Liao (2014), the models show there is a negative and significant association between loan loss provision with both ΔGDP and CSRET. Using the same reason, the models show there is a positive and significant association between the loan loss provision and the unemployment rate (UNEMP). Next, the results of the model show that the loan loss allowance (ALW_{t-1}) and the charge-off (CO) variables significantly associated with loan loss provision. As it also mentioned by Beatty and Liao (2014), there is a significant association between loan loss provision and both loan loss allowance and charge-off variables.

However, using different sample there are several differences between this thesis and Beatty and Liao (2014) results. First, in the model (c) and the model (d) the $SIZE_{t-1}$ variable become insignificant. However, the coefficient relatively the same. The next difference is in the model (d), the last year loan loss allowance (ALW_{t-1}) negatively associated with loan loss provision. This negative association appears after the net charge-off (CO) added in the model. The possible reason for the different results is this thesis uses more observations after the financial crisis. Therefore, in line with the counter pro-cyclical buffer policy, after the financial crisis banks become more conservative and tend to anticipate following year non-performing asset by increasing their loan loss allowance. Since this behavior can generate overestimate allowance, this can lead to lower loan loss provision in the current and future period. The comparative model from Bushman and Williams (2012) shows the same results. However, since Bushman and Williams

(2012) model uses the least independent variables compare to the other models, their model generates the smallest R-square and the highest residual[6].

5.3.1.2 Multicollinearity

Following Beatty and Liao (2014), this thesis also tests the loan loss provision model variables multicollinearity using Pearson correlation (Appendix 3a). In line with Beatty and Liao (2014) sample, there is a relatively high correlation between the macroeconomics variables, the change of gross domestic products (ΔGDP), the return of Case-Shiller index (CSRET), and the change in unemployment rate (ΔUNEMP). Since the loan loss allowance (ALW_{t-1}) and the charge-off (CO) capture the loan loss recognition, the two variables are also highly correlated.

5.3.2 Main regression analysis

This thesis then uses the residuals of each model in table 5 as the abnormal loan loss provision. In line with Beatty and Liao (2014) and Ma and Song (2016), this thesis uses each bank year average of abnormal loan loss provision as the proxy of earnings management.

5.3.2.1 Logit regression assumptions

Similar with the OLS regression, there are logit regression assumptions. The important assumptions are the independence of the error terms, the linearity of the independent variables, and little or zero multicollinearity. However, since there is no confirmed test to analyze the first two assumptions, this thesis tests the multicollinearity assumption. The Pearson correlation and the variance inflation factor (VIF) are presented in the Appendix 3b. The consensus is if the VIF value above 10 there is multicollinearity and therefore the assumption is violated. The results show that there are several variables with the VIF above 10 which are the minimum capital requirement (CAR), the moderating CAR variable (MCAR), and the bank size (SIZE). Furthermore, to mitigate the multicollinearity problem, this thesis tests the fraud incentives (hypotheses 2a) in different models separated with the moderating incentives (hypotheses 2b). In addition, this thesis also excludes the bank size from the control variables.

[6] The comparison of each model abnormal loan loss provision can be seen in table 15, appendix 3.

5.3.2.2 Regression analysis

5.3.2.2.1 Hypothesis 1 and Hypotheses 2a

Table 5 shows the logit regression (i.e. equation 8) to test the first hypothesis and hypotheses 2.a. The prior earnings management variables are from the five-loan loss provision accrual models in table 4.

Table 5 Logit regression

Financial statements fraud=$\alpha_0+\alpha_1$ARES_(a/b/c/d/e)+ α_2CC+α_3CAR+α_4B_DIS+α_5LIQ+α_6BIG4+α_7ROA+α_8ICMW+error

Variables	Prediction	Model (a) Coefficients (z-statistics)	Model (b) Coefficients (z-statistics)	Model (c) Coefficients (z-statistics)	Model (d) Coefficients (z-statistics)	Model (e) Coefficients (z-statistics)
Hypothesis 1						
ARES_a	+	132.549***				
		(6.67)				
ARES_b	+		135.053***			
			(6.17)			
ARES_c	+			188.454***		
				(6.12)		
ARES_d	+				181.334***	
					(5.91)	
ARES_e	+					150.258***
						(8.03)
Hypotheses 2a						
CC	+	182.074***	181.120***	180.440***	180.565***	179.447***
		(15.75)	(15.68)	(15.64)	(15.67)	(15.56)
CAR	-	-8.077***	-8.334***	-8.063***	-8.074***	-7.988***
		(-5.49)	(-5.67)	(-5.49)	(-5.49)	(-5.42)
B_DIS	-	-0.000**	-0.000**	-0.000**	-0.000**	-0.000**
		(-2.28)	(-2.29)	(-2.33)	(-2.38)	(-2.52)
LIQ	-	4.204***	4.300***	4.326***	4.309***	4.091***
		(5.36)	(5.51)	(5.54)	(5.52)	(5.19)
Control variables						
BIG4	-	0.720***	0.721***	0.725***	0.724***	0.705***
		(10.44)	(10.44)	(10.50)	(10.49)	(10.19)
ROA	-	-22.888***	-24.467***	-27.193***	-27.562***	-18.306**
		(-2.89)	(-3.10)	(-3.51)	(-3.55)	(-2.28)
ICMW	+	-0.021	-0.033	-0.064	-0.060	-0.048
		(-0.09)	(-0.14)	(-0.27)	(-0.25)	(-0.20)
constant	?	-2.499***	-2.461***	-2.485***	-2.479***	-2.495***
		(-11.58)	(-11.42)	(-11.54)	(-11.52)	(-11.57)
prob > chi-2		0.0000	0.0000	0.0000	0.0000	0.0000
pseudo r2		0.0995	0.0984	0.0983	0.0979	0.1027
N		9551	9551	9551	9551	9551

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Model (a)-model (d) are from Beatty and Liao (2014), model (e) is from Bushman and Williams (2012)

Definition of the variables:

ARES_a	Increase of abnormal loan loss provision model (a) from eight quarters before fraud period to fraud period.
ARES_b	Increase of abnormal loan loss provision model (b) from eight quarters before fraud period to fraud period.
ARES_c	Increase of abnormal loan loss provision model (c) from eight quarters before fraud period to fraud period.
ARES_d	Increase of abnormal loan loss provision model (d) from eight quarters before fraud period to fraud period.
ARES_e	Increase of abnormal loan loss provision model (e) from eight quarters before fraud period to fraud period.
CC	Cost of capital: interest expenses divided by total loans.
CAR	Tier-1 capital divided by weighted average asset risk (%).

B_DIS	Bank distress: DeLisle et al. (2007) Z-score.
LIQ	Liquidity variable: cash and equivalents divided by total assets.
BIG4	Dummy variable, 1 if the fiscal year auditor is the Big Four, 0 otherwise.
ROA	Income divided by average total assets.
ICMW	Dummy variable, 1 if there is an internal control material weakness, 0 otherwise.

The results in table 5 support the first hypothesis. In each model, the prior earnings management variable has a positive and significant association with financial statements fraud. It means banks with a higher increase in the abnormal loan loss provision in the last two years are more likely to commit financial statements fraud. This shows that prior earnings management can be used to predict financial statements fraud (Dechow et al., 2011: Perols & Lougee, 2011). The highest and the least coefficient of the prior earnings management variables are respectively from Beatty and Liao (2014) model (c) and model (a).

Regarding the first three of hypotheses 2.a, table 5 shows that in line with the hypotheses, banks are more likely to commit financial statements fraud if the banks have a higher cost of capital (hypothesis 2.1.a), lower minimum capital requirement (hypothesis 2.2.a), and higher financial distress (hypothesis 2.3.a). The results in line with the assumptions that earnings management incentives can also motivate firms to commit fraud (Dechow et al., 1996; Heally & Wahlen, 1999: Perols & Lougee, 2011). Regarding the cost of capital, in line with Dechow et al. (1996), banks with relatively higher interest expenses are more likely committing financial statements fraud to have lower external funding. In line with Heally and Wahlen (1999) and Ronen and Yaari (2008), banks are more likely committing financial statements fraud to avoid regulators scrutiny and other stakeholders' attention due to the lower minimum capital requirement. Regarding the bank distress, even though the variables coefficients in each model are slightly around zero, the results support the hypothesis that banks with higher failure risk are more likely committing financial statements fraud. The result is in line with Rosner (2003) who mentions that firms tend to manipulate their financial reports in the pre-bankruptcy period.

However, regarding the liquidity incentive, the results in table 5 shows that inconsistent with hypothesis 2.4.a, banks with higher liquidity are more likely committing financial statements fraud. The results in line with the descriptive analysis that shows there is no significant differences between the fraud group and the non-fraud group liquidity. In addition, the time-series analysis shows that before the fraud period, the fraud group liquidity is below the non-fraud group, nevertheless it gradually increases and then higher than the non-fraud group after the fraud period.

The possible explanation of the result is banks increase their liquidity in the fraud period to anticipate their stakeholders (e.g. creditors) negative reaction.

Regarding the control variables, table 5 shows that, in line with Beneish (1999a), Perols and Lougee (2011), and Uygur (2013), banks with lower profitability captured by return on assets (ROA) are more likely committing financial statements fraud. This is also consistent with the descriptive analysis that shows the profitability of the fraud group is lower than the non-fraud group. Next, inconsistent with the prediction, banks audited by the big four audit firms (BIG4) are more likely committing financial statements fraud. The results also show there is no significant association between internal control material weakness (ICMW) and financial statement fraud.

The statistics indicators in table 5 show that as the p-values of prob > chi-square in each model closely around zero, the independent variables are significantly associated with the dependent variable. Regarding the pseudo R-square, same with Beatty and Liao (2014), the explanatory power of the overall independent variables to the financial statements fraud relatively low. However, since this thesis examines the association between prior earnings management and financial statements fraud in the same picture with several fraud incentives, the pseudo R-squares in table 6 are higher than the pseudo R-square of Beatty and Liao (2014).

Fixed effect year

Table 6 shows the logit regressions with the year fixed effect variables. The fixed effect assumes that there is a time invariant factors in the variables tested in a regression. It means that the fixed effect variable can control the time invariant characteristics, so then the regression can measure the net effect of the independent variables to the dependent variables. Since there is a possibility of special events in particular period that may affect the outcome of the regression, this thesis controls the time fixed effect (i.e. fiscal year).

The results in table 6 show that the year fixed effect significantly impacts the outcome of the regressions. Inconsistent with the first hypothesis, prior earnings management has an insignificant association with financial statements fraud. The possible explanation of the results is the concentration of observations with higher fraud frequency and higher prior earnings management in the years around the financial crisis[7]. Without controlling the year fixed effect, the higher fraud frequency and the higher magnitude of prior earnings management in those years

[7] The fraud frequency distribution and the magnitude of abnormal loan loss provision in table 3.

dominate the same variables with the lower values in the other years. This explanation is consistent with the positive and significant coefficient of the fixed year effect from the year 2006-2010.

Table 6 Logit regression with year fixed effect

Financial statements fraud=α_0+α_1ARES_(a/b/c/d/e)+ α_2CC+α_3CAR+α_4B_DIS+α_5LIQ+α_6BIG4+α_7ROA+α_8ICMW+error

Variables	Prediction	Model (a) Coefficients (z-statistics)	Model (b) Coefficients (z-statistics)	Model (c) Coefficients (z-statistics)	Model (d) Coefficients (z-statistics)	Model (e) Coefficients (z-statistics)
Hypothesis 1						
ARES_a	+	-96.416^{***}				
		(-3.28)				
ARES_b	+		-99.022^{***}			
			(-3.05)			
ARES_c	+			-82.838^{*}		
				(-1.85)		
ARES_d	+				-92.513^{**}	
					(-2.09)	
ARES_e	+					-55.056^{**}
						(-2.00)
Hypotheses 2a						
CC	+	70.931	73.571^{*}	70.822	71.168	67.103
		(1.62)	(1.68)	(1.62)	(1.63)	(1.53)
CAR	-	-1.746	-1.718	-1.915	-1.947	-1.737
		(-0.64)	(-0.63)	(-0.71)	(-0.72)	(-0.64)
B_DIS	-	0.000	0.000	0.000	0.000	0.000
		(0.31)	(0.28)	(0.23)	(0.24)	(0.31)
LIQ	-	2.768^{*}	2.644^{*}	2.692^{*}	2.701^{*}	2.751^{*}
		(1.76)	(1.69)	(1.73)	(1.73)	(1.76)
Control variables						
BIG4	-	0.368	0.375	0.388	0.389	0.374
		(1.52)	(1.55)	(1.61)	(1.61)	(1.54)
ROA	-	-17.877	-17.573	-15.074	-15.379	-16.037
		(-1.62)	(-1.59)	(-1.37)	(-1.39)	(-1.45)
ICMW	+	0.551^{*}	0.555^{*}	0.506	0.519	0.514
		(1.65)	(1.66)	(1.52)	(1.56)	(1.54)
Year fixed effect						
Fiscal Year=2003		0.000	0.000	0.000	0.000	0.000
		(.)	(.)	(.)	(.)	(.)
Fiscal Year=2004		0.078	0.057	0.106	0.107	0.104
		(0.18)	(0.13)	(0.25)	(0.25)	(0.24)
Fiscal Year=2005		-0.367	-0.379	-0.365	-0.365	-0.346
		(-0.84)	(-0.87)	(-0.83)	(-0.83)	(-0.79)
Fiscal Year=2006		1.198^{***}	1.190^{***}	1.215^{***}	1.214^{***}	1.280^{***}
		(2.90)	(2.90)	(2.95)	(2.95)	(3.10)
Fiscal Year=2007		2.241^{***}	2.209^{***}	2.216^{***}	2.217^{***}	2.305^{***}
		(5.31)	(5.26)	(5.27)	(5.27)	(5.43)
Fiscal Year=2008		3.583^{***}	3.551^{***}	3.511^{***}	3.520^{***}	3.569^{***}
		(9.22)	(9.21)	(9.09)	(9.10)	(9.13)
Fiscal Year=2009		2.429^{***}	2.387^{***}	2.292^{***}	2.306^{***}	2.358^{***}
		(6.29)	(6.25)	(6.01)	(6.04)	(6.08)
Fiscal Year=2010		1.058^{***}	1.009^{**}	1.004^{**}	1.007^{**}	1.082^{***}
		(2.64)	(2.53)	(2.51)	(2.52)	(2.68)
Fiscal Year=2011		-0.375	-0.320	-0.239	-0.246	-0.273
		(-0.81)	(-0.70)	(-0.52)	(-0.54)	(-0.60)
Fiscal Year=2012		0.458	0.515	0.530	0.531	0.483
		(1.06)	(1.20)	(1.23)	(1.23)	(1.12)
Fiscal Year=2013		-0.721	-0.742	-0.751	-0.748	-0.701

	(-1.40)	(-1.44)	(-1.46)	(-1.45)	(-1.36)
Fiscal Year=2014	0.715^{*}	0.697	0.722^{*}	0.724^{*}	0.779^{*}
	(1.65)	(1.62)	(1.67)	(1.68)	(1.80)
Fiscal Year=2015	1.116^{***}	1.098^{***}	1.125^{***}	1.128^{***}	1.187^{***}
	(2.64)	(2.62)	(2.67)	(2.68)	(2.81)
Prob > chi2	0.0000	0.0000	0.0000	0.0000	0.0000
LR chi2(20)	1140.43	1138.96	1133.12	1134.03	1133.70
N	6839	6839	6839	6839	6839

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Model (a)-model (d) are from Beatty and Liao (2014), model (e) is from Bushman and Williams (2012)

Definition of the variables:

ARES_a	Increase of abnormal loan loss provision model (a) from eight quarters before fraud period to fraud period.
ARES_b	Increase of abnormal loan loss provision model (b) from eight quarters before fraud period to fraud period.
ARES_c	Increase of abnormal loan loss provision model (c) from eight quarters before fraud period to fraud period.
ARES_d	Increase of abnormal loan loss provision model (d) from eight quarters before fraud period to fraud period.
ARES_e	Increase of abnormal loan loss provision model (e) from eight quarters before fraud period to fraud period.
CC	Cost of capital: interest expenses divided by total loans.
CAR	Tier-1 capital divided by weighted average asset risk (%).
B_DIS	Bank distress: DeLisle et al. (2007) Z-score.
LIQ	Liquidity variable: cash and equivalents divided by total assets.
BIG4	Dummy variable, 1 if the fiscal year auditor is the Big Four, 0 otherwise.
ROA	Income divided by average total assets.
ICMW	Dummy variable, 1 if there is an internal control material weakness, 0 otherwise.

Regarding the fraud incentives, the results in table 6 inconsistent with hypotheses 2.a. After controlling the year fixed effect, there is an insignificant association between a cost of capital (CC), minimum capital requirement (CAR), and bank distress (B_DIS) with financial statements fraud. However, inconsistent with the other models and in line with hypothesis 2.1.a, in the model (b) there is a positive and significant association between a cost of capital and financial statements fraud. Regarding the liquidity variable, in all models, the liquidity variable is positively associated financial statements fraud. The result is in line with the result in table 5 and therefore inconsistent with hypothesis 2.4.a.

Regarding the control variables, inconsistent with the prediction, big four audit firms (BIG4) and banks profitability as it captured by return on assets (ROA) have an insignificant association with financial statements fraud. However, the results of the model (a) and the model (b) show that banks with internal control material weakness (ICMW) are more likely to commit financial statements fraud.

5.3.2.2.2 Hypotheses 2b

Table 7 shows the logit regression with the moderating variables (i.e. equation 9). Equation 9 is particularly used to test hypotheses 2.b.

Table 7 Logit regression (hypotheses 2b)

Financial statements fraud=$\alpha_0+\alpha_1$ARES_(a/b/c/d/e)+ α_2CC+α_3CAR+α_4B_DIS+α_5LIQ+α_6MCC(a/b/c/d/e)+α_7MCAR(a/b/c/d/e)+α_8MB_DIS(a/b/c/d/e)+α_9MLIQ(a/b/c/d/e) +α_{10}BIG4+α_{11}ROA+α_{12}ICMW+error

Variables	Prediction	Model (a) Coefficients (z-statistics)	Model (b) Coefficients (z-statistics)	Model (c) Coefficients (z-statistics)	Model (d) Coefficients (z-statistics)	Model (e) Coefficients (z-statistics)
Hypothesis 1						
ARES_a	+	-19.880				
		(-0.17)				
ARES_b	+		42.431			
			(0.33)			
ARES_c	+			72.540		
				(0.40)		
ARES_d	+				59.444	
					(0.33)	
ARES_e	+					138.284
						(1.32)
Hypotheses 2a						
CC	+	180.229***	180.040***	180.173***	180.040***	177.468***
		(15.44)	(15.54)	(15.53)	(15.51)	(15.11)
CAR	-	-8.951***	-8.681***	-8.719***	-8.734***	-8.593***
		(-5.86)	(-5.75)	(-5.73)	(-5.74)	(-5.62)
B_DIS	-	-0.000**	-0.000**	-0.000**	-0.000**	-0.001***
		(-2.33)	(-2.07)	(-2.27)	(-2.35)	(-3.90)
LIQ	-	4.647***	4.444***	4.461***	4.458***	4.493***
		(5.87)	(5.64)	(5.68)	(5.67)	(5.64)
Hypotheses 2b						
MCC_(a/b/c/d/e)	+	3786.704	2005.360	-13563.554	-11928.921	-3605.127
		(0.49)	(0.24)	(-1.18)	(-1.04)	(-0.49)
MCAR_(a/b/c/d/e)	+	704.698	416.475	1440.607	1359.055	-210.094
		(0.97)	(0.51)	(1.17)	(1.11)	(-0.32)
MBDIS_(a/b/c/d/e)	+	0.772***	0.414**	0.560**	0.609**	0.998***
		(4.91)	(2.46)	(2.27)	(2.45)	(6.14)
MLIQ_(a/b/c/d/e)	+	-391.419	-161.549	-495.406	-481.765	-643.307
		(-0.90)	(-0.33)	(-0.69)	(-0.67)	(-1.62)
Control variables						
BIG4	-	0.734***	0.723***	0.727***	0.726***	0.695***
		(10.60)	(10.47)	(10.53)	(10.51)	(10.02)
ROA	-	-31.468***	-29.477***	-32.368***	-32.912***	-26.500***
		(-3.91)	(-3.65)	(-4.10)	(-4.16)	(-3.29)
ICMW	+	-0.061	-0.036	-0.045	-0.045	-0.096
		(-0.26)	(-0.15)	(-0.19)	(-0.19)	(-0.41)
constant	?	-2.384***	-2.411***	-2.393***	-2.384***	-2.362***
		(-10.75)	(-10.94)	(-10.80)	(-10.76)	(-10.57)
Prob > chi-2		0.0000	0.0000	0.0000	0.0000	0.0000
Pseudo r2		0.1056	0.1018	0.1035	0.1032	0.1154
N		9551	9551	9551	9551	9551

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Model (a)-model (d) are from Beatty and Liao (2014), model (e) is from Bushman and Williams (2012)

Definition of the variables:

ARES_a Increase of abnormal loan loss provision model (a) from eight quarters before fraud period to fraud period.
ARES_b Increase of abnormal loan loss provision model (b) from eight quarters before fraud period to fraud period.
ARES_c Increase of abnormal loan loss provision model (c) from eight quarters before fraud period to fraud period.
ARES_d Increase of abnormal loan loss provision model (d) from eight quarters before fraud period to fraud period.
ARES_e Increase of abnormal loan loss provision model (e) from eight quarters before fraud period to fraud period.

CC	Cost of capital: interest expenses divided by total loans.
CAR	Tier 1 capital divided by weighted average asset risk (%).
B_DIS	Bank distress: DeLisle et al. (2007) Z-score.
LIQ	Liquidity variable: cash and equivalents divided by total assets.
MCC	Moderating variable model (a/b/c/d/e) abnormal loan loss provision with cost of capital (interest expense divided by total liabilities).
MCAR	Moderating variable model (a/b/c/d/e) abnormal loan loss provision with tier 1 capital (%).
MB_DIS	Moderating variable model (a/b/c/d/e) abnormal loan loss provision with bank distress (DeLisle et al. (2007) Z-score).
MLIQ	Moderating variable model (a/b/c/d/e) abnormal loan loss provision with liquidity (cash & equivalents divided by total assets).
BIG4	Dummy variable, 1 if the fiscal year auditor is the Big Four, 0 otherwise.
ROA	Income divided by average total assets.
ICMW	Dummy variable, 1 if there is an internal control material weakness, 0 otherwise.

Hypotheses 2.b state that the fraud incentives can strengthen the association between the prior earnings management with the financial statements fraud. However, the results do not in line with the hypotheses. Despite the results that the moderating variable of bank distress (MBDIS) has a positive and significant association with financial statements fraud, the results do not completely support hypothesis 2.3.b, since in the same regression, there is no significant association between prior earnings management and financial statements fraud. The other moderating variables of fraud incentives which are a cost of capital (MCC), minimum capital requirement (MCAR), and liquidity (MLIQ) have an insignificant association with financial statements fraud.

The results in table 7 inconsistent with the first hypothesis, the association between prior earnings management and financial statements fraud become insignificant in each model. This is inconsistent with the results in table 5. The different results can be caused by the multicollinearity in table 7 or the correlated omitted variables in table 5. However, regarding the hypotheses 2.a, table 7 have the same results with the table 5. The results consistent with hypothesis 2.1.a, hypothesis 2.1.b, and hypothesis 2.1.c, and inconsistent with hypothesis 2.1.d.

Regarding the control variables, table 7 have the same results with table 6. Consistent with the prediction, bank profitability has a negative and significant association with financial statements fraud. Inconsistent with the prediction, bank audited by the big four audit firms are more likely committing financial statements fraud. Finally, there is no significant association between internal control material weakness and financial statements fraud.

Year fixed effect

Table 8 shows equation 8 with year fixed effect. Regarding hypotheses 2.b, table 8 has the same results with table 7 and therefore the results do not support hypotheses 2.b. With the year fixed effect, the moderating variable of bank distress (MBDIS) have a positive and significant association with financial statements fraud. However, since there is no significant association between prior earnings management and financial statements fraud, the results do not fully support hypothesis 2.2.c. The results of the other moderating variables have the same results with the results in table 7 and therefore do no support the hypotheses.

Table 8 Logit regression with year fixed effect (hypotheses 2b)

Financial statements fraud=$\alpha_0+\alpha_1$ARES_(a/b/c/d/e)+ α_2CC+α_3CAR+α_4B_DIS+α_5LIQ+α_6MCC(a/b/c/d/e)+α_7MCAR(a/b/c/d/e)+α_8MB_DIS(a/b/c/d/e)+α_9MLIQ(a/b/c/d/e) +α_{10}BIG4+α_{11}ROA+α_{12}ICMW+error

Variables	Prediction	Model (a) Coefficients (z-statistics)	Model (b) Coefficients (z-statistics)	Model (c) Coefficients (z-statistics)	Model (d) Coefficients (z-statistics)	Model (e) Coefficients (z-statistics)
Hypothesis 1						
ARES_a	+	-277.193*				
		(-1.79)				
ARES_b	+		-327.813*			
			(-1.92)			
ARES_c	+			-247.216		
				(-1.00)		
ARES_d	+				-257.316	
					(-1.04)	
ARES_e	+					20.717
						(0.14)
Hypotheses 2a						
CC	+	67.788	78.162*	75.217*	74.645*	68.320
		(1.54)	(1.78)	(1.71)	(1.70)	(1.55)
CAR	-	-2.201	-1.745	-2.295	-2.304	-1.997
		(-0.79)	(-0.63)	(-0.83)	(-0.83)	(-0.72)
B_DIS	-	0.000	0.000	0.000	0.000	-0.000
		(0.34)	(0.20)	(0.18)	(0.20)	(-0.03)
LIQ	-	2.870*	2.461	2.599	2.562	2.930*
		(1.79)	(1.53)	(1.61)	(1.58)	(1.82)
Hypotheses 2b						
MCCa/b/c/d/e	+	-1933.649	-3207.688	-22488.003	-18776.354	-15088.429
		(-0.17)	(-0.27)	(-1.33)	(-1.11)	(-1.41)
MCARa/b/c/d/e	+	1274.175	1581.584	1596.403	1411.291	-324.766
		(1.27)	(1.43)	(0.96)	(0.85)	(-0.36)
MBDISa/b/c/d/e	+	0.365*	0.120	0.561*	0.571*	0.484**
		(1.86)	(0.58)	(1.89)	(1.93)	(2.50)
MLIQa/b/c/d/e	+	39.969	863.011	609.522	682.988	-254.021
		(-1.18)	(0.58)	(0.55)	(0.62)	(-1.76)
Control variables						
BIG4	-	0.371	0.360	0.374	0.375	0.375
		(1.53)	(1.48)	(1.53)	(1.54)	(1.55)
ROA	-	-23.401**	-20.261*	-19.326*	-19.577*	-20.285*
		(-2.07)	(-1.79)	(-1.72)	(-1.74)	(-1.82)
ICMW	+	0.564*	0.567*	0.512	0.523	0.499
		(1.70)	(1.72)	(1.56)	(1.58)	(1.50)
Year fixed effect						
Fiscal Year=2003		0.000	0.000	0.000	0.000	0.000
		(.)	(.)	(.)	(.)	(.)
Fiscal Year=2004		0.092	0.070	0.101	0.103	0.118
		(0.21)	(0.16)	(0.23)	(0.24)	(0.27)
Fiscal Year=2005		-0.343	-0.376	-0.360	-0.357	-0.335
		(-0.78)	(-0.86)	(-0.82)	(-0.81)	(-0.76)
Fiscal Year=2006		1.241***	1.185***	1.206***	1.211***	1.217***
		(2.99)	(2.87)	(2.92)	(2.93)	(2.95)
Fiscal Year=2007		2.257***	2.214***	2.200***	2.203***	2.221***
		(5.34)	(5.27)	(5.22)	(5.22)	(5.24)
Fiscal Year=2008		3.585***	3.605***	3.545***	3.546***	3.478***
		(9.18)	(9.30)	(9.12)	(9.12)	(8.85)

Fiscal Year=2009	2.438***	2.414***	2.300***	2.314***	2.291***
	(6.31)	(6.30)	(6.00)	(6.04)	(5.91)
Fiscal Year=2010	1.068***	1.026**	1.026**	1.025**	1.026**
	(2.66)	(2.56)	(2.55)	(2.55)	(2.54)
Fiscal Year=2011	-0.364	-0.266	-0.197	-0.209	-0.313
	(-0.78)	(-0.58)	(-0.43)	(-0.45)	(-0.67)
Fiscal Year=2012	0.464	0.554	0.571	0.566	0.470
	(1.06)	(1.28)	(1.32)	(1.31)	(1.07)
Fiscal Year=2013	-0.731	-0.695	-0.689	-0.696	-0.817
	(-1.41)	(-1.35)	(-1.33)	(-1.34)	(-1.57)
Fiscal Year=2014	0.756*	0.781*	0.817*	0.804*	0.701
	(1.73)	(1.80)	(1.87)	(1.84)	(1.61)
Fiscal Year=2015	1.150***	1.180***	1.213***	1.202***	1.087**
	(2.70)	(2.78)	(2.85)	(2.83)	(2.55)
Prob > chi2	0.0000	0.0000	0.0000	0.0000	0.0000
LR chi2(20)	1145.81	1143.85	1140.15	1140.37	1141.16
N	6839	6839	6839	6839	6839

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Model (a)-model (d) are from Beatty and Liao (2014), model (e) is from Bushman and Williams (2012)

Definition of the variables:

ARES_a	Increase of abnormal loan loss provision model (a) from eight quarters before fraud period to fraud period.
ARES_b	Increase of abnormal loan loss provision model (b) from eight quarters before fraud period to fraud period.
ARES_c	Increase of abnormal loan loss provision model (c) from eight quarters before fraud period to fraud period.
ARES_d	Increase of abnormal loan loss provision model (d) from eight quarters before fraud period to fraud period.
ARES_e	Increase of abnormal loan loss provision model (e) from eight quarters before fraud period to fraud period.
CC	Cost of capital: interest expenses divided by total loans.
CAR	Tier 1 capital divided by weighted average asset risk (%).
B_DIS	Bank distress: DeLisle et al. (2007) Z-score.
LIQ	Liquidity variable: cash and equivalents divided by total assets.
MCC	Moderating variable model (a/b/c/d/e) abnormal loan loss provision with cost of capital (interest expense divided by total liabilities).
MCAR	Moderating variable model (a/b/c/d/e) abnormal loan loss provision with tier 1 capital (%).
MB_DIS	Moderating variable model (a/b/c/d/e) abnormal loan loss provision with bank distress (DeLisle et al. (2007) Z-score).
MLIQ	Moderating variable model (a/b/c/d/e) abnormal loan loss provision with liquidity (cash & equivalents divided by total assets).
BIG4	Dummy variable, 1 if the fiscal year auditor is the Big Four, 0 otherwise.
ROA	Income divided by average total assets.
ICMW	Dummy variable, 1 if there is an internal control material weakness, 0 otherwise.

The results in table 8 inconsistent with the first hypothesis. There is there an insignificant association between prior earnings management and financial statements fraud in the model (c), model (d), and model (e), and there is a negative and significant association in the model (a) and model (b). Regarding hypotheses 2.a, there is a positive and significant association between a cost of capital (CC) and financial statements fraud in the model (b), model (c), and model (d). However, there is no significant association between minimum capital requirement (CAR) and bank distress with financial statements fraud. Regarding the liquidity incentive, inconsistent with hypothesis 2.4.a, the results show there is a positive and significant association between bank liquidity and financial statements fraud in the model (a) and model (e).

Regarding the control variables, in line the results in table 7, the results in table 8 show there is an insignificant association between big four audit firms (BIG4) with financial statements fraud. Consistent with the prediction, banks with lower profitability are more likely committing financial statements fraud. Finally, there is a positive and significant association between internal control material weakness (ICMW) in the model (a) and model (b). However, there is an insignificant association in the model (c), model (d), and model (e).

5.3 Additional tests

5.4.1 Alternative proxies of the earnings management variable

5.4.1.1 Abnormal loan loss provision

5.4.1.1.1 Beatty and Liao (2014) current abnormal loan loss provision

This thesis different to Beatty and Liao (2014) who use current abnormal loan loss provision to predict provision manipulation captured by restatements and SEC comment letters. Following Perols and Lougee (2011) who use prior discretionary accrual to predict financial statements fraud, this thesis uses the increase of abnormal loan loss provision in the last three years to capture banks prior earnings management. In Appendix 5, using Beatty and Liao (2014) current abnormal loan loss provision approach, this thesis has the same results. However, the coefficient of earnings management and the R-square of the models using the prior earnings management is higher than the models using the current abnormal loan loss provision. It also means that using the sample in this thesis, the prediction power of the model using prior abnormal loan loss provision is higher than the model using current abnormal loan loss provision. Therefore in the prediction model, together with the fraud incentives variables, this thesis uses the prior earnings management variables to predict financial statements fraud.

5.4.1.1.2 Bushman and Williams (2012) loan loss provision model

This thesis uses Bushman and Williams (2012) loan loss provision accrual model (i.e. model (e)) as the comparison model to the Beatty and Liao (2014) models. In general, model (e) results in line with the Beatty and Liao (2014) models. However, because the aim of this paper is to find the opportunistic loan loss provision, following Beatty and Liao (2014), this thesis excludes two independent variables which are earnings before loan loss provision (EBP) and tier-1 capital requirement (CAR). Ma and Song (2016) mention that earnings before loan loss provision is used in the loan loss provision accrual model to identify managers' earnings smoothing behavior. In Appendix 4, this thesis also uses Bushman and Williams (2012) loan loss provision original model.

First, Appendix 6, table 22 shows that model (e) results consistent with Bushman and Williams (2012) original model (i.e. model f). In addition, the earnings before loan loss provision has a negative and significant association with loan loss provision. It means that banks use loan loss provision to smooth their earnings. However, there is no significant association between minimum capital requirement and loan loss provision. Second, regarding the hypotheses test, using abnormal loan loss provision from the model (f), the results in table 23 show there is an insignificant association between prior earnings management with financial statements fraud. However, without the year fixed effect, the results in line with hypothesis 2.1.a and hypothesis 2.2.a.

5.4.2 Delayed loan loss recognition

Since delayed loan loss recognition can also capture bank opportunity behavior (Bushman & Williams, 2015), this thesis also tests the association between financial statements fraud and delayed loan loss recognition. Following Beatty and Liao (2011) and Bushman and Williams (2015) delayed loan loss recognition approaches, to run the regressions in equation 4, this thesis requires minimum 12 observations of each bank. Therefore, the sample reduces from 9,715 to 6,936 observations. Following the steps in section 4.2.2.2, the values of the R-square differences between equation 4 model (1) and equation 4 model (2) generate the dummy variable (DELAY). The dummy variable (DELAY) equal to 1 if banks have a greater delay on loan loss recognition, and equal to 0 if banks have a small delay on loan loss recognition.

Table 9 shows two regressions respectively from equation 8 and equation 9. To capture current and prior earnings management there are two delayed loan loss recognition variables in table 9. First, current delayed loan loss recognition (DELAY). Second, delayed loan loss recognition four quarters before (DELAY_4). The results in table 9 show that banks with a higher delay on loan loss recognition have a positive and significant association with financial statements fraud. However, different to abnormal loan loss provision, the results still remain with the year fixed effect. Regarding the hypotheses 2.b. different to abnormal loan loss provision, without the year fixed effect, the cost of capital strengthens the association between delay loan loss recognition and financial statements fraud.

Table 9 Logit regression with delayed loan loss recognition

Regression 1:
Financial statements fraud=$\alpha_0+\alpha_1$DELAY/4+ α_2CC+α_3CAR+α_4B_DIS+α_5LIQ +α_6BIG4+α_7ROA+α_8ICMW+error

Regression 2:
Financial statements fraud=$\alpha_0+\alpha_1$DELAY/4+ α_2CC+α_3CAR+α_4B_DIS+α_5LIQ+α_6MCC_D+α_7MCAR_D+α_8MB_DIS_D+α_9MLIQ_D +α_{10}BIG4+α_{11}ROA+α_{12}ICMW+error

	Prediction	Regression 1		Regression 2		Regression 1		Regression 2	
		random effect b/z	fixed effect b/z	random effect b/z	fixed effect b/z	random effect b/z	fixed effect b/z	random effect b/z	fixed effect b/z
Hypothesis 1									
DELAY	+	0.433***	0.320**	-0.125	0.274				
		(5.69)	(3.23)	(-0.27)	(0.46)				
DEL_4	+					0.294***	0.194	0.176	0.356
						(3.81)	(1.96)	(0.37)	(0.59)
Hypothesis 2a									
CC	+	181.335***	26.012	155.157***	23.251	180.947***	27.216	161.369***	21.469
		(14.40)	(0.53)	(9.20)	(0.46)	(14.21)	(0.55)	(9.99)	(0.42)
T1CAP	-	-13.005***	-4.968	-13.957***	-5.252	-13.009***	-5.008	-12.576***	-4.239
		(-7.66)	(-1.61)	(-6.06)	(-1.45)	(-7.68)	(-1.63)	(-5.55)	(-1.16)
B_DIS	-	-0.000*	0.000	-0.001*	0.000	-0.000*	0.000	-0.000	0.000
		(-2.24)	(0.44)	(-2.37)	(0.12)	(-2.35)	(0.63)	(-1.19)	(0.92)
LIQ	-	3.489***	2.616	4.312***	3.727	3.327***	2.513	3.308**	3.049
		(4.10)	(1.54)	(4.16)	(1.86)	(3.89)	(1.48)	(3.10)	(1.42)
Hypothesis 2b									
CC_D/4	+			59.400*	12.056			50.863	22.528
				(2.35)	(0.35)			(1.96)	(0.63)
CAR_D/4	+			2.121	0.331			-0.757	-1.560
				(0.63)	(0.08)			(-0.23)	(-0.38)
BDIS_D/4	+			0.000	0.000			-0.000	-0.000
				(1.27)	(0.55)			(-0.79)	(-0.65)
LIQ_D/4	+			-2.491	-2.426			-0.142	-0.832
				(-1.39)	(-1.08)			(-0.08)	(-0.37)
Control variables									
BIG4	-	0.610***	0.292	0.612***	0.293	0.614***	0.265	0.613***	0.264
		(7.83)	(0.96)	(7.84)	(0.97)	(7.89)	(0.87)	(7.87)	(0.87)
ROA	-	-27.324***	-11.301	-27.355***	-11.940	-29.053***	-12.229	-28.806***	-12.202
		(-3.40)	(-0.95)	(-3.40)	(-1.01)	(-3.64)	(-1.05)	(-3.60)	(-1.05)
ICMW	+	-0.164	-0.003	-0.206	-0.011	-0.123	0.016	-0.099	0.035
		(-0.57)	(-0.01)	(-0.70)	(-0.03)	(-0.43)	(0.04)	(-0.35)	(0.09)
Fiscal Year=2006			0.000		0.000		0.000		0.000
			(.)		(.)		(.)		(.)
Fiscal Year=2007			0.786***		0.750***		0.991***		1.018***
			(4.57)		(4.26)		(5.75)		(5.70)
Fiscal Year=2008			1.941***		1.913***		2.093***		2.103***
			(10.09)		(9.83)		(11.20)		(11.18)
Fiscal Year=2009			0.607*		0.589*		0.766**		0.783**
			(2.26)		(2.19)		(2.92)		(2.96)
Fiscal Year=2010			-0.577		-0.592		-0.422		-0.400
			(-1.72)		(-1.76)		(-1.28)		(-1.20)
Fiscal Year=2011			-1.806***		-1.811***		-1.641***		-1.614***
			(-4.31)		(-4.32)		(-3.97)		(-3.87)
Fiscal Year=2012			-1.135**		-1.131**		-0.978*		-0.940*
			(-2.81)		(-2.79)		(-2.45)		(-2.34)
Fiscal Year=2013			-2.645***		-2.643***		-2.483***		-2.450***
			(-4.99)		(-4.97)		(-4.71)		(-4.63)
Fiscal Year=2014			-0.999*		-1.005*		-0.848*		-0.814
			(-2.39)		(-2.39)		(-2.05)		(-1.95)
Fiscal Year=2015			-0.486		-0.499		-0.347		-0.306
			(-1.20)		(-1.23)		(-0.87)		(-0.76)
constant		-1.783***		-1.516***		-1.702***		-1.662***	
		(-7.08)		(-4.61)		(-6.77)		(-5.12)	
R-sqr		0.1078		0.1101		0.1047		0.1058	
LR chi2(21)			853.25		855.21		846.81		848.42
N		6927	4979	6927	4979	6923	4975	6923	4975

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Definition of the variables:

DELAY	Dummy variable equal to 1 if the bank has higher delay on loan loss recognition, and equal to 0 if the bank has lower delay on loan loss recognition
DEL_4	Dummy variable equal to 1 if the bank has higher delay on loan loss recognition, and equal to 0 if the bank has lower delay on loan loss recognition
CC	Cost of capital: interest expenses divided by total loans
CAR	Tier 1 capital divided by weighted average asset risk (%)
B_DIS	Bank distress: DeLisle et al. (2007) Z-score
LIQ	Liquidity variable: cash and equivalents divided by total assets
MCC_D/4	Moderating variable model (/4) delayed loan loss provision with cost of capital (interest expense divided by total liabilities)
MCAR_D/4	Moderating variable model (/4) delayed loan loss provision with tier 1 capital (%)
MBDIS_D/4	Moderating variable model (/4) delayed loan loss provision with bank distress (DeLisle et al. (2007) Z-score)
MLIQ_D/4	Moderating variable model (/4) delayed loan loss provision with liquidity (cash & equivalents divided by total assets)
BIG4	Dummy variable, 1 if the fiscal year auditor is the Big Four, 0 otherwise
ROA	Income divided by average total assets
ICMW	Dummy variable, 1 if there is an internal control material weakness, 0 otherwise

5.4.3 Alternative proxies of the fraud incentives variables

5.4.3.1 Bank distress

In addition to DeLisle et al. (2007) Z-score, following Chiaramonte et al. (2016), this thesis also uses Z-score from Boyd and Graham et al. (1986) and Boyd et al. (2007). The results in Appendix 7, table 25 and table 26 show that the main results still hold.

5.4.3.2 Liquidity

In addition to the liquidity variable in the main analysis, this thesis also uses net stable funding ratio (NSFR) as the proxy of liquidity. Following Chiaramonte and Casu (2016) this thesis uses similar items available in Bank Compustat to calculate NSFR. However, since NSFR used specific items, using NSFR as the liquidity variable decreases the observations. The results in Appendix 7, table 27 and table 28 show that the main results still hold.

5.4 Prediction models

Following Dechow et al. (2011), this thesis uses backward elimination method and excludes the insignificant variables from the models. In addition, the full prediction model in this thesis excludes the internal control material weakness (ICMW) variable which has an insignificant association with financial statements fraud. In addition, in line with Beneish (1999a) and Dechow et al., (2011), this thesis also excludes the big four audit firms (BIG4) from the prediction models.

First, this thesis only uses the abnormal loan loss provision and the fraud incentives separately to predict financial statements fraud. Table 10 shows the prediction models using the abnormal loan loss provision from Beatty and Liao (2014) model (a) - model (e) and the fraud incentives. Following Beatty and Liao (2014), to measure the models' prediction accuracy, this thesis classifies the models' results as a fraud observation if the result is more than 50% and otherwise. The results in table 10 show that prediction accuracy of the both regressions relatively low. The prediction results show that the models difficult to predict fraud observations and have a high type-2 error. Since the R-square of regression 2 (i.e. only with fraud incentives) higher than regression 1 (i.e. only with prior earnings management), regression 2 can predict fraud observations better than regression 1.

Table 10 The prediction models

Regression 1: financial statements fraud=$\alpha_0+\alpha_1$ARES(a/b/c/d/e) +error

Regression 2: financial statements fraud=$\alpha_0+\alpha_1$CC+α_2CAR+α_3B_DIS+α_4LIQ+α_5ROA+error

Variables	Prediction	Regression 1					Regression 2
		Model (a)	Model (b)	Model (c)	Model (d)	Model (e)	Incentives
		Coefficients (z-statistics)	Coefficients (z-statistics)	Coefficients (z-statistics)	Coefficients (z-statistics)	Coefficients (z-statistics)	Coefficients (z-statistics)
ARES_a	+	206.185*** (11.66)					
ARES_b	+		210.703*** (10.92)				
ARES_c	+			314.898*** (11.38)			
ARES_d	+				310.208*** (11.23)		
ARES_e	+					218.556*** (13.18)	
CC	+						180.785*** (15.90)
T1CAP	-						-10.035*** (-6.98)
B_DIS	-						-0.000** (-2.81)
LIQ							4.207*** (5.40)
ROA							-34.394*** (-4.66)
constant	?	-2.136*** (-63.39)	-2.130*** (-63.47)	-2.131*** (-63.55)	-2.130*** (-63.56)	-2.151*** (-63.15)	-1.832*** (-9.03)
pseudo r2		0.0198	0.0173	0.0186	0.0181	0.0257	0.0757
N		9707	9707	9707	9707	9707	9559
The prediction results based on Beatty and Liao (2014) approach							
True positive		0.00%	0.00%	0.00%	0.00%	0.00%	1.40%
True negative		99.90%	99.91%	99.88%	99.88%	99.90%	98.28%
Type-1 error		0.10%	0.09%	0.12%	0.12%	0.10%	1.72%
Type-2 error		100.00%	100.00%	100.00%	100.00%	100.00%	98.60%

Table 11 shows the prediction models using both prior earnings management and the fraud incentives variables. The results in table 11 are relatively the same with the results in table 10, regression 2. The full prediction models are more likely to predict non-fraud observations and less likely to predict fraud observations. This issue is caused by using the random sample to develop the fraud prediction models (Beneish, 1999a)[8].

[8] Therefore, Benesih (1999)a does not use random sample and uses matched manipulators and non-manipulators firms.

Table 11 The full prediction models

Financial statements fraud=$\alpha_0+\alpha_1$ARES(a/b/c/d/e)+ α_2CC+α_3CAR+α_4B_DIS+α_5LIQ +α_6BIG4+α_7ROA+error						
Variables	Prediction	Model (a) Coefficients (z-statistics)	Model (b) Coefficients (z-statistics)	Model (c) Coefficients (z-statistics)	Model (d) Coefficients (z-statistics)	Model (e) Coefficients (z-statistics)
ARES_a	+	134.695***				
		(6.77)				
ARES_b	+		134.847***			
			(6.25)			
ARES_c	+			182.937***		
				(6.07)		
ARES_d	+				176.910***	
					(5.88)	
ARES_e	+					157.828***
						(8.43)
CC	+	177.642***	176.635***	176.072***	176.219***	174.731***
		(15.41)	(15.34)	(15.33)	(15.35)	(15.18)
CAR	-	-9.294***	-9.580***	-9.331***	-9.335***	-9.179***
		(-6.42)	(-6.63)	(-6.46)	(-6.47)	(-6.32)
B_DIS	-	-0.000**	-0.000**	-0.000**	-0.000**	-0.000**
		(-2.01)	(-2.02)	(-2.06)	(-2.10)	(-2.16)
LIQ	-	4.087***	4.190***	4.222***	4.206***	3.965***
		(5.20)	(5.35)	(5.40)	(5.38)	(5.02)
ROA	-	-17.287**	-18.943**	-21.829***	-22.165***	-11.941
		(-2.20)	(-2.42)	(-2.83)	(-2.88)	(-1.49)
constant	?	-1.989***	-1.946***	-1.965***	-1.961***	-2.001***
		(-9.65)	(-9.47)	(-9.58)	(-9.56)	(-9.70)
pseudo r2		0.0828	0.0817	0.0813	0.0810	0.0867
N		9551	9551	9551	9551	9551
The prediction results based on Beatty and Liao (2014) approach						
True positive		1.30%	1.21%	1.30%	1.30%	1.30%
True negative		98.21%	98.21%	98.19%	98.18%	98.22%
Type-1 error		1.79%	1.79%	1.81%	1.82%	1.78%
Type-2 error		98.70%	98.79%	98.70%	98.70%	98.70%

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Model (a)-model (d) are from Beatty and Liao (2014), model (e) is from Bushman and Williams (2012)

Definition of the variables:

ARES_a	Increase of abnormal loan loss provision model (a) from eight quarters before fraud period to fraud period.
ARES_b	Increase of abnormal loan loss provision model (b) from eight quarters before fraud period to fraud period.
ARES_c	Increase of abnormal loan loss provision model (c) from eight quarters before fraud period to fraud period.
ARES_d	Increase of abnormal loan loss provision model (d) from eight quarters before fraud period to fraud period.
ARES_e	Increase of abnormal loan loss provision model (e) from eight quarters before fraud period to fraud period.
CC	Cost of capital: interest expenses divided by total loans.
CAR	Tier 1 capital divided by weighted average asset risk (%).
B_DIS	Bank distress: DeLisle et al. (2007) Z-score.
LIQ	Liquidity variable: cash and equivalents divided by total assets.
ROA	Income divided by average total assets.

In addition to Beatty and Liao (2014) approach, since Dechow et al. (2011) also use a random sample of firms with accounting misstatements, this thesis also use Dechow et al. (2011) approach to classify the prediction results. First, Dechow et al. (2011) calculate the unconditional probability that the prediction model predicts fraud observation. The unconditional probability is generated from the prediction models using the mean values of the independent variables. Next, to get the F-score value, the predicted values of each observation are then compared with the unconditional probability (Appendix 9).

Table 12 shows the prediction results of Dechow et al. (2011) approach. The results show that Dechow et al. (2013) approach can predict fraud observations better than Beatty and Liao (2014). It also shows that the lower the cut-off score, the higher the accuracy of fraud prediction and the type-1 error.

Table 12 Prediction results with cut-off score approach

	Model (a)	Model (b)	Model (c)	Model (d)	Model (e)
Cut-off score =1					
True positive	77.26%	77.35%	77.07%	76.98%	78.47%
True negative	55.11%	55.75%	55.69%	55.65%	56.02%
Type-1 error	44.89%	44.25%	44.31%	44.35%	43.98%
Type-2 error	22.74%	22.65%	22.93%	23.02%	21.53%
Cut-off score =2					
True positive	39.79%	38.12%	38.49%	38.21%	40.91%
True negative	84.08%	84.14%	84.66%	84.68%	83.48%
Type-1 error	15.92%	15.86%	15.34%	15.32%	16.52%
Type-2 error	60.21%	61.88%	61.51%	61.79%	59.09%

Model (a)-model (d) are from Beatty and Liao (2014), model (e) is from Bushman and Williams (2012)

Out of sample test

This thesis also tests the prediction model to the test observations. The test observations are 2.738 out of sample observations that contain 131 fraud observations and 2.607 non-fraud observations. The test observations are the bank-quarter observations from the main observations that cannot be merged with the other datasets. The prediction results in table 13 show that the prediction models can predict fraud observations better in the test observations.

Table 13 Out of the sample prediction results

	Model (a)	Model (b)	Model (c)	Model (d)	Model (e)
Beatty and Liao (2014) approach					
True positive	11.45%	12.21%	12.98%	12.98%	12.21%
True negative	93.17%	93.06%	93.17%	93.33%	93.17%
Type-1 error	6.83%	6.94%	6.83%	6.67%	6.83%
Type-2 error	88.55%	87.79%	87.02%	87.02%	87.79%
Dechow et al. (2011) approach with cut-off score 1					
True positive	78.63%	79.39%	80.15%	82.44%	78.63%
True negative	51.32%	51.75%	51.75%	51.40%	53.01%
Type-1 error	48.68%	48.25%	48.25%	48.60%	46.99%
Type-2 error	21.37%	20.61%	19.85%	17.56%	21.37%

Model (a)-model (d) are from Beatty and Liao (2014), model (e) is from Bushman and Williams (2012)

6. Conclusion

Financial statements fraud causes damages to firms' stakeholders. The damages become more severe to the financial stability if the opportunistic accounting policy is from the banking industry (Bushman & Williams, 2015; Ma & Song, 2015). Prior literature examines financial statements fraud to predict and to prevent financial statements fraud in the future. However, there are limited studies of financial statements fraud in the banking industry. Using the first assumption that firms commonly engaged with earnings management before committing financial statements fraud (Beneish, 1999a; Dechow et al., 2011; Perols & Lougee, 2011), and the second assumption that fraud incentives can motivate firms to commit accounting manipulation (Dechow et al., 1996, Heally & Wahlen, 1999, Ronen & Yaari, 2008), this thesis finds the answer to the following research question:

"Do prior earning management and fraud incentives increase the likelihood of financial statements fraud in the banking industry?

Perols and Lougee (2011) find that prior earnings management can predict financial statements fraud. In the banking industry, Beatty and Liao (2014) find that abnormal loan loss positively associated with financial statements fraud. Hence, the first hypothesis:

H.1: Banks with prior abnormal loan loss provision are more likely to commit financial statements fraud.

The results do not completely support the first hypothesis. Table 6 shows that prior earnings management as it captured by the increasing of abnormal loan loss provision in the last three years has a positive and significant association with financial statements fraud. However, using the year fixed effect, table 7 shows that prior earnings management has an insignificant association with financial statements fraud. The possible explanation of the result is the concentration of both the higher fraud frequency and the higher magnitude of abnormal loan loss provision around the financial crisis years. In the additional test, current abnormal loan loss provision from Beatty and Liao (2014) and delayed loan loss recognition from Bushman and Williams (2015) has a positive and significant association with financial statements fraud with and without year fixed effect.

The results have several implications. First, as the loan loss provision model of Beatty and Liao (2014) and Bushman and Williams (2011) can capture the macroeconomic condition around the financial crisis years, the results show that the loan loss provision quality decrease during the financial crisis years. Second, the association between earnings management and financial statements fraud is stronger around the financial crisis years. Therefore, since both financial statements fraud and earnings management decrease bank transparency, the results of this thesis in line with Bushman and Williams (2015) and Ma and Song (2016) who find that opportunistic earnings management as it captured respectively by delayed loan loss provision and abnormal loan loss provision has a positive association with bank opacity and bank risk. Therefore, the results support the future implication of expected loan loss provision as the part of the counter cyclical policy. Third, since there is an increase of abnormal loan loss provision before the fraud period, the results also suggest that banks also use real earnings management despite (e.g. loans restructuring) before performing earnings management through loan loss provision.

The second hypotheses assume there is an association between fraud incentives circumstance with financial statements in the banking industry. The hypotheses based on Dechow et al. (1996), Heally and Wahlen (1999), and Ronen and Yaari (2008) findings that several earnings management incentives can also motivate financial statements fraud. Since firms manage their earnings to achieve a low cost of capital (Dechow et al., 1996; Stolowy & Breton, 2004), the hypotheses are:

H.2.1.a: Banks with a higher cost of capital are more likely to commit financial statements fraud.

H.2.1.b: Banks with prior earnings management are more likely to commit financial statements fraud the more they have higher interest expenses.

Using interest expenses divided by total liabilities as the proxy of the cost of capital. The results (e.g. model b) show that consistent with hypothesis 2.1.a, banks with a higher cost of capital (i.e. cost of debt) are more likely to commit financial statements fraud. However, as the moderating variable, inconsistent with hypothesis 2.1.b, the moderating variable of the cost of capital has an insignificant association with financial statements fraud. The inconsistent results of hypothesis 2.1.b can be caused by the positive association between the abnormal loan loss provision and interest expenses. It means banks with higher interest expenses are reluctant to increase their loan loss provision since it can give more cost to their earnings.

Heally and Wahlen (1999) and Ronen and Yaari (2008) mention that minimum capital requirement in the banking industry can motivate earning management. Hence the hypotheses:

H.2.2.a: Banks with lower minimum capital requirement are more likely to commit financial statements fraud.

H.2.2.b: Banks with prior earnings management are more likely to commit financial statements fraud the more they have a lower minimum capital requirement.

The results do not completely in line with hypothesis 2.2.a and hypothesis 2.2.b. The results show that there is a negative and significant association between the minimum capital requirement and financial statements fraud. However, the minimum capital requirement association becomes insignificant after considering the year fixed effect. Regarding the moderating variable, the association between the minimum capital requirement and financial statements fraud is insignificant. The insignificant association of minimum capital requirement as the moderating variable of abnormal loan loss provision can be explained by the exclusion of the loan loss provision from the tier-1 capital requirement (Ahmed et al., 1999; Beatty & Liao, 2014).

Next, since firms tend to manipulate their financial statements before going to failure (Rosner, 2003) and firms with extreme financial characteristics are more likely manipulate their earnings (Beneish, 1997), the hypotheses:

H.2.3.a: Banks with higher financially distress are more likely to commit financial statements fraud.

H.2.3.b: Banks with prior earnings management are more likely to commit financial statements fraud the more they faced higher distress.

In line with hypothesis 2.3.a, banks with higher financial distress are more likely to commit financial statements fraud. However, with the year fixed effect, the association becomes insignificant. The possible explanation of the inconsistency results after year fixed effect is the higher fraud frequency of fraud and bank failure are concentrated around financial crisis years[9], thus it is difficult to find the association outside the period. Regarding bank distress as a moderating variable, the results do not support hypothesis 2.3.b. In the additional tests the results still the same.

Since banks with lower liquidity are more likely to fail (Chiaramonte et al., 2016), and firms with less liquidity tend to manipulate their earnings Beneish (1997:1999a), the hypotheses:

H.2.4.a: Banks with lower liquidity are more likely to commit financial statements fraud.

H.2.4.b: Banks with prior earnings management are more likely to commit financial statements fraud the more they have lower liquidity.

[9] Table 1. Chiaramonte et al. (2016)

The results are inconsistent with the both hypotheses. First, banks with higher liquidity tend to commit financial statements fraud. The results still hold with net stable funding ratio (NSFR) as the additional liquidity variable. Second, there is an insignificant association between liquidity as moderating variable with financial statements fraud. Using the time-series analysis, it can be seen that before the fraud period, the fraud group liquidity is lower than the non-fraud group. However, after the fraud period, the fraud group liquidity is higher than the non-fraud group. Since the observations of this thesis include the financial crisis years, the results can be caused by banks' policy to withhold the loans growth or to anticipate the future liquidity problems (Beatty and Liao, 2011).

Based on the association analysis above, this thesis develops financial statements fraud prediction models. Two approaches to get the prediction model results are from Dechow et al. (2011) and Beatty and Liao (2014). The prediction models are then tested to out of sample observations, using Dechow et al. (2011) the prediction models have type-1 and type-2 errors respectively 72.49% and 11.15%[10].

Contributions

This thesis contributes to accounting manipulation and banking literature. In the earnings management literature, following Beatty and Liao (2014) future research suggestion, this thesis examines earnings management through abnormal loan loss provision and delayed loan loss recognition in the same picture with the fraud incentives. The results show that the fraud incentives significantly associated with the financial statements fraud. However, the fraud incentives do not strengthen nor weakens the association between abnormal loan loss provision and financial statements fraud. Next, since current literature of financial statement fraud do not include the financial crisis years in their observations (e.g. Dechow et al. 2011; Perols & Lougee, 2011), commonly exclude the banking industry (e.g. Beneish 1999a; Perols & Lougee, 2011), this thesis also contributes to managers' opportunistic behavior in the financial crisis and in the banking industry.

In the banking industry, in line with Bushman and Williams (2015), Ma and Song (2016), the results of this thesis also suggest that there is an increase of bank opacity in financial crisis

[10] The result is from the prediction model (a) i.e. using abnormal loan loss provision from Beatty and Liao (2014) model (a).

years as it captured by the higher frequency of financial statements fraud and the higher magnitude of abnormal loan loss provision.

For the practitioners, in addition to the existing financial statements fraud prediction model (e.g. Beneish, 1999a; Dechow et al., 2011; Beatty and Liao 2014), the banking-specific fraud prediction model in this thesis can also help auditors and regulators to identify banks with higher financial statements fraud risk. Since this thesis incorporates bank regulations (i.e. minimum capital requirement) and regulators restatements, this thesis also contributes to banking regulation enforcement study.

Limitations and future research

First, following Dechow et al. (2011) and Perols and Lougee (2011) this thesis uses specific accounting policy (i.e. loan loss provision and delayed loan loss recognition) to predict various financial statements fraud in the banking industry. The future research can examine a specific type of financial statements fraud in the banking industry. Second, this thesis suggests that banks also performing real earnings management before committing financial statements fraud. However, this thesis does not examine the existence of loan loss restructuring around the financial statements fraud period. Third, the fraud prediction models in this thesis is developed from the sample that including the financial crisis years. Therefore, there is a risk of overfitting or the models cannot work the other observations. To mitigate the overfitting risk, following Perols et al. (2016), future research can focus to develop a prediction model with data analytics methods. Finally, since the prediction model in this thesis uses future data (i.e. next year non-performing assets), the models cannot predict future financial statements fraud. However, the models are more benefit to classify firms with a high risk of financial statements fraud.

References

Ahmed, A., Takeda, C., & 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. doi:10.1016/S0165-4101(99)00017-8

Anandarajan, A., Hasan, I., & McCarthy, C. (2007). Use of loan loss provisions for capital, earnings management and signalling by australian banks. *Accounting and Finance, 47*(3), 357.

Armstrong, C., Jagolinzer, A., & Larcker, D. (2010). Chief executive officer equity incentives and accounting irregularities. *Journal of Accounting Research, 48*(2), 225-271. doi:10.1111/j.1475-679X.2009.00361.x

Armstrong, C. S., Larcker, D. F., Ormazabal, G., & Taylor, D. J. (2013). The relation between equity incentives and misreporting: The role of risk-taking incentives. *Journal of Financial Economics, 109*(2), 327.

Association of Certified Fraud Examiners. (2016). *Report to the nations on occupational fraud and abuse 2016 global fraud study*. Austin: ACFE Inc.

Aydemir, R., & Guloglu, B. (2017). How do banks determine their spreads under credit and liquidity risks during business cycles? *Journal of International Financial Markets, Institutions & Money,46*, 147-157. doi:10.1016/j.intfin.2016.08.001

Bischof, Jannis and Daske, Holger and Elfers, Ferdinand and Hail, Luzi, A Tale of Two Regulators: Risk Disclosures, Liquidity, and Enforcement in the Banking Sector (February 27, 2016). Available at SSRN: https://ssrn.com/abstract=2580569

Beatty, A. L., Ke, B., & Petroni, K. R. (2002). Earnings management to avoid earnings declines across publicly and privately held banks. *The Accounting Review, 77*(3), 547-570. doi:10.2308/accr.2002.77.3.547

Beatty, A., & Liao, S. (2011). Do delays in expected loss recognition affect banks' willingness to lend? *Journal of Accounting & Economics, 52*(1), 1-1.

Beatty, A., & Liao, S. (2014). Financial accounting in the banking industry: A review of the empirical literature. *Journal of Accounting & Economics, 58*(2-3), 339.

Beaver, W., & Engel, E. (1996). Discretionary behavior with respect to allowances for loan losses and the behavior of security prices. *Journal of Accounting & Economics, 22*(1-3), 177-206.

Beneish, M. D. (1997). Detecting GAAP violation: Implications for assessing earnings management among firms with extreme financial performance. *Journal of Accounting and Public Policy, 16*(3), 271-309.

Beneish, M. D. (1999a). The detection of earnings manipulation. *Financial Analysts Journal, 55*(5), 24-36.

Beneish, M. D. (1999b). Incentives and penalties related to earnings overstatements that violate GAAP. *The Accounting Review, 74*(4), 425-457.

Beneish, M. (2001). Earnings management: A perspective. *Managerial Finance, 27*(12), 3-17.

Betz, F., Oprică, S., Peltonen, T. A., & Sarlin, P. (2014). Predicting distress in European banks. *Journal Of Banking & Finance, 45*(2), 225-241. doi:10.1016/j.jbankfin.2013.11.041

Bornemann, S., Kick, T., Memmel, C., & Pfingsten, A. (2012). Are banks using hidden reserves to beat earnings benchmarks? evidence from germany. *Journal of Banking and Finance, 36*(8), 2403-2415.

Boyd, J., & Graham, S. (1986). Risk, regulation, and bank holding company expansion into nonbanking. *Federal Reserve Bank of Minneapolis. Quarterly Review - Federal Reserve Bank of Minneapolis,*

Boyd, J., & De Nicolo, G. (2005). The theory of bank risk taking and competition revisited. *The Journal of Finance, 60*(3), 1329-1343.

Bushman, R. M. (2014). Thoughts on financial accounting and the banking industry. *Journal of Accounting & Economics, 58*(2-3), 384.

Bushman, R. M., & Williams, C. D. (2012). Accounting discretion, loan loss provisioning, and discipline of banks' risk-taking. *Journal of Accounting & Economics, 54*(1), 1.

Bushman, R. M., & Williams, C. D. (2015). Delayed expected loss recognition and the risk profile of banks. *Journal of Accounting Research, 53*(3).

Chiaramonte, L., & Casu, B. (2016). *Capital and liquidity ratios and financial distress. evidence from the european banking industry*. London: Academic Press.

Chiaramonte, L., Liu, (. H., Poli, F., & Zhou, M. (2016). How accurately can z-score predict bank failure? *Financial Markets, Institutions & Instruments, 25*(5), 333-360. doi:10.1111/fmii.12077

Costello, A. M., Granja, J., & Weber, J. (2016). Do Strict Regulators Increase the Transparency of the Banking System?.

Craig Nichols, D., Wahlen, J., & Wieland, M. (2009). Publicly traded versus privately held: Implications for conditional conservatism in bank accounting. *Review of Accounting Studies, 14*(1), 88-122. doi:10.1007/s11142-008-9082-3

DeAngelo, H., DeAngelo, L., & Skinner, D. (2004). Are dividends disappearing? dividend concentration and the consolidation of earnings. *Journal of Financial Economics, 72*(3).

Dechow, P. (1994). Accounting earnings and cash flows as measures of firm performance: The role of accounting accruals. *Journal of Accounting & Economics, 18*(1), 3-3.

Dechow, P. M., & Dichev, I. D. (2002). The quality of accruals and earnings: The role of accrual estimation errors. *The Accounting Review, 77*, 35-59.

Dechow, P., Ge, W., & 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. doi:10.1016/j.jacceco.2010.09.001

Dechow, P. M., Ge, W., Larson, C. R., & Sloan, R. G. (2011). Predicting material accounting misstatements. *Contemporary Accounting Research, 28*(1), 17-82. doi:10.1111/j.1911-3846.2010.01041.x

Dechow, P., & Skinner, D. (2000). Earnings management: Reconciling the views of accounting academics, practitioners, and regulators. *Accounting Horizons, 14*(2), 235-250.

Dechow, P. M., Sloan, R. G., & Sweeney, A. P. (1996). Causes and consequences of earnings manipulation: An analysis of firms subject to enforcement actions by the SEC*. *Contemporary Accounting Research, 13*(1), 1-36.

Desai, H., Rajgopal, S., & Yu, J. J. (2016). Were information intermediaries sensitive to the financial statement-based leading indicators of bank distress prior to the financial crisis? *Contemporary Accounting Research, 33*(2), 576-606.

Dichev, I. D., Graham, J. R., Harvey, C. R., & Rajgopal, S. (2013). Earnings quality: Evidence from the field. *Journal of Accounting and Economics, 56*, 1-33. Retrieved from http://hdl.handle.net/2324/1124701

Doyle, J., Ge, W., & McVay, S. (2007). Accruals quality and internal control over financial reporting. *The Accounting Review, 82*(5), 1141-1170.

Ellul, A., & Yerramilli, V. (2013). Stronger risk controls, lower risk: Evidence from U.S. bank holding companies. *The Journal of Finance, 68*(5), 1757.

Erickson, M., Hanlon, M., & Maydew, E. (2006). Is there a link between executive equity incentives and accounting fraud? *Journal of Accounting Research, 44*(1), 113-143. doi:10.1111/j.1475-679X.2006.00194.x

Fanning, K., & Cogger, K. (1998). Neural network detection of management fraud using published financial data. *Intelligent Systems in Accounting, Finance & Management, 7*(1), 21-41. doi:10.1002/(SICI)1099-1174(199803)7:13.0.CO;2-K

Graham, J., Harvey, C., & Rajgopal, S. (2005). The economic implications of corporate financial reporting. *Journal of Accounting and Economics, 40*(1-3), 3-73. doi:10.1016/j.jacceco.2005.01.002

Gompers, P., Ishii, J., & Metrick, A. (2003). Corporate governance and equity prices. *The Quarterly Journal of Economics, 118*(1), 107-155.

Healy, P. M., & Wahlen, J. M. (1999). A review of the earnings management literature and its implications for standard setting. *Accounting Horizons, 13*(4), 365-383.

Jones, J. (1991). Earnings management during import relief investigations. *Journal of Accounting Research, 29*(2).

Jones, K. L., Krishnan, G. V., & Melendrez, K. D. (2008). Do models of discretionary accruals detect actual cases of fraudulent and restated earnings? an empirical analysis. *Contemporary Accounting Research, 25*(2), 499-531. doi:10.1506/car.25.2.8

Kim, D., & Sohn, W. (2017). The effect of bank capital on lending: Does liquidity matter? *Journal of Banking & Finance, 77*.

Kothari, S., Leone, A., & Wasley, C. (2005). Performance matched discretionary accrual measures. *Journal of Accounting & Economics, 39*(1), 163-197.

Laeven, L., & Levine, R. (2009). Bank governance, regulation and risk taking. *Journal of Financial Economics, 93*(2), 259.

Ma, C. (1988). Loan loss reserves and income smoothing: The experience in the u.S. banking industry. *Journal of Business Finance & Accounting, 15*(4), 487-487.

Ma, M., & Song, V. (2016). Discretionary loan loss provisions and systemic risk in the banking industry. *Accounting Perspectives, 15*(2).

Perols, J. L., Bowen, R. M., Zimmermann, C., & Samba, B. (2016). Finding needles in a haystack: Using data analytics to improve fraud prediction. *The Accounting Review,*

Perols, J. L., & Lougee, B. A. (2011). The relation between earnings management and financial statement fraud. *Advances in Accounting, Incorporating Advances in International Accounting, 27*(1), 39-53.

Picker, R., Leo, K. J., Loftus, J., Wise, V. P., Clark, K. C., & Alfredson, K. (2013). *Applying international financial reporting standards* (3rd edition. ed.). Milton, Qld : Wiley,.

Price, R. A., Sharp, N. Y., & Wood, D. A. (2011). Detecting and predicting accounting irregularities: A comparison of commercial and academic risk measures. *Accounting Horizons, 25*(4), 755-780. doi:10.2308/acch-50064

Richardson, S., Sloan, R., Soliman, M., & Tuna, I. (2006). The implications of accounting distortions and growth for accruals and profitability. *The Accounting Review, 81*(3), 713-743.

Ronen, J., & Yaari, V. (2008). *Earnings management : Emerging insights in theory, practice, and research*. New York : Springer,.

Rosner, Rebecca L. "Earnings manipulation in failing firms." *Contemporary Accounting Research* 20.2 (2003): 361-408.

Stolowy, Hervé, and Gaétan Breton. "Accounts manipulation: A literature review and proposed conceptual framework." *Review of Accounting and Finance* 3.1 (2004): 5-92.

Strobl, Günter. "Earnings manipulation and the cost of capital." *Journal of Accounting Research* 51.2 (2013): 449-473.

Uygur, O. (2013). Earnings management and executive compensation. *Banking and Finance Review, 5*(2), 33-54. Retrieved from http://www.econis.eu/PPNSET?PPN=821049453

Watts, R., & Zimmerman, J. (1986). *Positive accounting theory* (Prentice hall contemporary topics in accounting series). Englewood Cliffs, N.J.: Prentice-Hall International.

Worrell, Mr DeLisle, Andrea M. Maechler, and Ms Srobona Mitra. *Decomposing financial risks and vulnerabilities in Eastern Europe*. No. 7-248. International Monetary Fund, 2007.

Appendix 1 – Summary of important literature

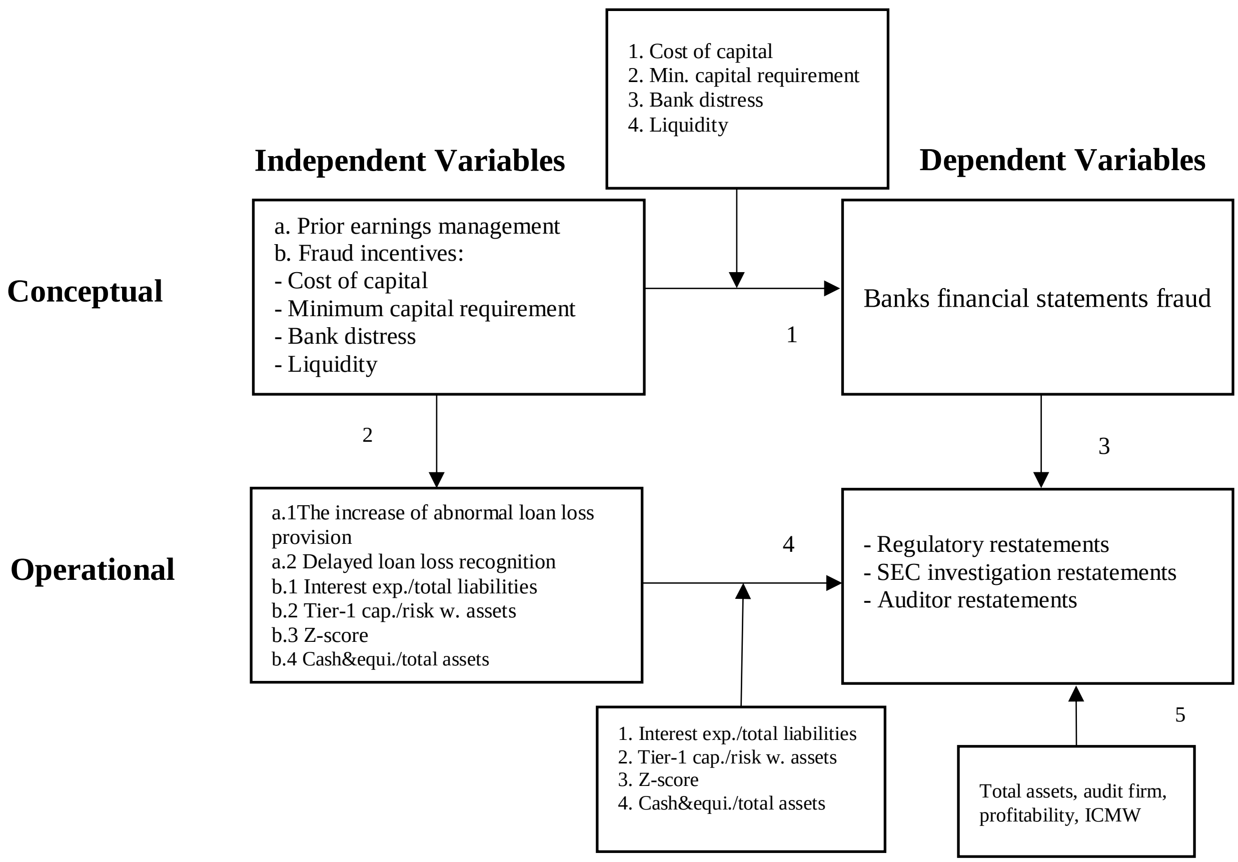
Table 14 Summary of the important literature

No.	Literature	Observations	Summary
1.	Dechow et al. (1996)	1982-1992 Various industries excluding the banking industry	- This paper is used to develop the hypotheses and the research design. - The dependent variable is from the SEC AAERs. - Match the manipulators group with the non-manipulators group. - This paper uses logit regression and finds that several incentives that can motivate earnings manipulation (e.g. lower cost of capital).
2.	Beneish (1999a)	1982-1992 Various industries excluding the banking industry	- This paper is used to develop the hypotheses. - The earnings manipulations, the dependent variable is from SEC AAERs and accounting restatements. - Matching the manipulators group with the non-manipulators group. - This paper uses probit regression and develops M-score model to predict earnings manipulations.
3.	Dechow et al. (2011)	1982-2005 Various industries including the banking industry	- This paper is used to develop the hypotheses and the research design. - The accounting misstatements, the dependent variable is from SEC AAERs. - Randomly assigned the misstating the firms and the non-misstating firms. - This paper uses time-series analysis to examine the characteristic of misstating firms. - This papers uses logit regression and develops F-score model to predict accounting misstatements.
4.	Beatty and Liao (2014)	1993-2012 From Compustat Banks	- This paper is used to develop the hypotheses and the research design. - This paper develops loan loss provision accrual model that is used in this thesis. - This papers uses restatements and SEC comments letter as the proxy of loan loss provision manipulation. - Randomly assigned the manipulation firms and the non-manipulation firms. - This paper uses logit regression to find the association between abnormal loan loss provision and loan loss provision manipulation.
5.	Bushman and Williams (2015)	1993-2009 From Compustat Bank and Bank Call Report	- This paper is used to develop the research design. - This papers uses delay loan loss provision recognition as the proxy of bank earnings management.

Appendix 2 – Libby boxes

Figure 3

Libby boxes



Appendix 3 - Time-series analysis of the variables

Table 15 Non-performing assets and loan loss provision changes

Variable	LLP Model R-square	Full sample mean (1)	Fraud mean (2)	Non-Fraud mean (3)	Difference (3) - (2)	P-value
			four quarters after fraud quarter			
NPA4		0.0220	0.0282	0.0213	-0.0069	0.0000
llp4		0.0018	0.0034	0.0016	-0.0018	0.0000
ARESa4	0.3147	0.0014	0.0021	0.0013	-0.0007	0.0000
ARESb4	0.4053	0.0013	0.0018	0.0012	-0.0006	0.0000
ARESc4	0.7396	0.0008	0.0011	0.0007	-0.0004	0.0000
ARESd4	0.7411	0.0008	0.0011	0.0007	-0.0004	0.0000
ARESe4	0.2319	0.0015	0.0022	0.0014	-0.0008	0.0000
			fraud quarter			
NPA		0.0220	0.0212	0.0221	0.0009	0.2749
llp		0.0018	0.0027	0.0017	-0.0011	0.0000
ARESa	0.3147	0.0014	0.0017	0.0014	-0.0003	0.0000
ARESb	0.4053	0.0013	0.0015	0.0013	-0.0002	0.0000
ARESc	0.7396	0.0008	0.0009	0.0007	-0.0002	0.0000
ARESd	0.7411	0.0008	0.0009	0.0007	-0.0002	0.0000
ARESe	0.2319	0.0015	0.0019	0.0014	-0.0004	0.0000
			four quarters before fraud quarter			
NPA_4		0.0220	0.0137	0.0231	0.0094	0.0000
llp_4		0.0018	0.0017	0.0018	0.0001	0.3526
ARESa_4	0.3147	0.0014	0.0013	0.0014	0.0002	0.0001
ARESb_4	0.4053	0.0013	0.0011	0.0013	0.0002	0.0000
ARESc_4	0.7396	0.0008	0.0007	0.0008	0.0001	0.0028
ARESd_4	0.7411	0.0008	0.0007	0.0008	0.0001	0.0063
ARESe_4	0.2319	0.0015	0.0014	0.0015	0.0001	0.1352
			eight quarters before fraud quarter			
NPA_8		0.0221	0.0123	0.0233	0.0109	0.0000
llp_8		0.0018	0.0013	0.0018	0.0005	0.0000
ARESa_8	0.3147	0.0014	0.0011	0.0015	0.0004	0.0000
ARESb_8	0.4053	0.0013	0.0010	0.0013	0.0003	0.0000
ARESc_8	0.7396	0.0008	0.0006	0.0008	0.0002	0.0000
ARESd_8	0.7411	0.0008	0.0006	0.0008	0.0002	0.0000
ARESe_8	0.2319	0.0015	0.0012	0.0015	0.0004	0.0000
N		9715	1073	8642		

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$. Model (a)-model (d) are from Beatty and Liao (2014), model (e) is from Bushman and Williams (2012)

Definition of the variables:

NPa	non-performing assets divided by total loans.
LLP	total loan loss provision divided by total loans.
ARESa	average of absolute abnormal loan loss provision model (a)
ARESb	average of absolute abnormal loan loss provision model (b)
ARESc	average of absolute abnormal loan loss provision model (c)
ARESd	average of absolute abnormal loan loss provision model (d)
ARESe	average of absolute abnormal loan loss provision model (e)

Table 16 Cost of capital, CAR, bank distress, and liquidity changes

	Full sample mean (1)	Fraud mean (2)	Non-Fraud mean (3)	Difference (3) - (2)	P-value
			four quarters after fraud period		
CC4	0.0043	0.0049	0.0043	-0.0007	0.0000
CAR4	0.0043	0.0061	0.0041	-0.0020	0.0000
BDIS4	0.0043	0.0065	0.0041	-0.0025	0.0000
LIQ4	0.0043	0.0059	0.0041	-0.0018	0.0000
			fraud period		
CC	12.2630	11.6119	12.3439	0.7319	0.0000
CAR	12.2611	11.1215	12.4026	1.2811	0.0000
B_DIS	12.2608	11.0488	12.4114	1.3627	0.0000
LIQ	12.2612	11.0972	12.4059	1.3087	0.0000
			four quarters before fraud period		
CC_4	231.6083	156.2416	241.0103	84.7687	0.0000
CAR_4	231.6451	264.0886	227.6031	-36.4855	0.0000
BDIS_4	231.6852	270.5391	226.8330	-43.7061	0.0000
LIQ_4	231.6180	203.9502	235.0753	31.1251	0.0000
			eight quarters before fraud period		
CC_8	0.0415	0.0464	0.0408	-0.0056	0.0000
CAR_8	0.0415	0.0413	0.0415	0.0001	0.9157
BDIS_8	0.0415	0.0375	0.0419	0.0044	0.0008
LIQ_8	0.0415	0.0367	0.0421	0.0053	0.0000
N	9715	1073	8642		

Definition of the variables:

CC	Cost of capital: interest expenses divided by total loans
CAR	Tier 1 capital divided by weighted average asset risk
B_DIS	Bank distress: DeLisle et al. (2007) Z-score.
LIQ	Liquidity: cash and equivalents divided by total assets

Appendix 4a – Multicollinearity loan loss provision model variables

Table 17 Pearson correlation of loan loss provision model variables

	LLP	ΔNPA t+1	ΔNPA t	ΔNPA t-1	ΔNPA t-2	SIZE	ΔLoan	ΔGDP	CSRET	ΔUNEMP	ALW_1	CO
LLP	1											
ΔNPA t+1	0.108***	1										
ΔNPA	0.187***	-0.142***	1									
ΔNPA t-1	0.157***	0.0942***	-0.181***	1								
ΔNPA t-2	0.191***	0.0751***	0.0933***	-0.210***	1							
SIZE_1	0.101***	-0.0207**	-0.0188*	-0.0138	-0.0145	1						
ΔLoan	-0.266***	0.0331***	0.0109	-0.0302***	-0.0477***	-0.0450***	1					
ΔGDP	-0.410***	-0.168***	-0.178***	-0.151***	-0.160***	-0.0447***	0.209***	1				
CSRET	-0.329***	0.0347***	0.00295	-0.0148	-0.0358***	-0.00267	0.235***	0.289***	1			
ΔUNEMP	0.296***	0.167***	0.179***	0.134***	0.131***	-0.0153	-0.125***	-0.653***	-0.0951***	1		
ALW_1	0.397***	-0.111***	-0.0741***	-0.0343***	-0.0284***	0.120***	-0.288***	-0.0887***	-0.414***	-0.00848	1	
CO	0.818***	0.00612	0.0435***	0.0904***	0.125***	0.135***	-0.324***	-0.296***	-0.371***	0.170***	0.582***	1

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Definition of the variables:

LLP	the loan loss provision divided by lagged total loans.
ΔNPA	the change of non-performing assets divided by lagged total loans.
SIZE	the natural log of total assets.
ΔLoan	the change of total loan divided by lagged total loans.
ΔGDP	the change of Gross Domestic Product over the quarter.
CSRET	the return of the Case-Shiller Real Estate Index over the quarter.
ΔUNEMP	the change of unemployment rates over the quarter.
ALW	the loan loss allowance divided by total loan.
CO	the net charge-off divided by lagged total loan.

Appendix 4b – Multicollinearity logit regression variables

Table 18 Pearson correlation of logit regression variables

	FSF	ARES_a	CC	T1CAP	B_DIS	LIQ	SIZE	ROA	BIG4	ICMW
FSF	1									
ARES_a	0.120***	1								
CC	0.206***	0.148***	1							
T1CAP	-0.138***	-0.142***	-0.371***	1						
B_DIS	-0.0439***	-0.0867***	-0.00699	-0.00383	1					
LIQ	0.000272	-0.0116	-0.249***	0.144***	-0.0837***	1				
SIZE	0.171***	0.0130	-0.0782***	-0.180***	0.0185	0.0963***	1			
ROA	-0.0889***	-0.302***	-0.0633***	0.119***	0.222***	-0.0445***	0.0318**	1		
BIG4	0.107***	-0.0122	-0.0272**	-0.104***	0.0574***	-0.0216*	0.532***	0.0330**	1	
ICMW	0.0107	0.0622***	0.0151	-0.0112	-0.0634***	-0.00941	-0.00866	-0.0390***	0.00156	1

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

ARES_a Increase of abnormal loan loss provision model (a) from eight quarters before fraud period to fraud period
CC Cost of capital: interest expenses divided by total loans
CAR Tier 1 capital divided by weighted average asset risk (%)
B_DIS Bank distress: DeLisle et al. (2007) Z-score.
LIQ Liquidity variable: cash and equivalents divided by total assets
SIZE Natural log of total assets
BIG4 Dummy variable, 1 if the fiscal year auditor is the Big Four, 0 otherwise
ROA Income divided by average total assets
ICMW Dummy variable, 1 if there is an internal control material weakness, 0 otherwise

Table 19 VIF test

Model (a)		Model (b)		Model (c)		Model (d)		Model (e)	
Variable	VIF	Variable	VIF	Variable	VIF	Variable	VIF	Variable	VIF
ARES_a	36.47	ARES_b	37.86	ARES_c	38.17	ARES_d	38.25	ARES_e	35.8
CARa	21.16	CARb	22.78	CARc	23.33	CARd	23.26	CARe	21.29
SIZE	18.32	SIZE	18.31	SIZE	18.26	SIZE	18.26	SIZE	18.22
CAR	11.63	CAR	11.58	CAR	11.66	CAR	11.67	CAR	11.67
MCC	4.98	MCC	4.87	MCC	5.24	MCC	5.27	MCC	4.82
CC	2.92	CC	2.92	CC	2.92	CC	2.92	CC	2.91
MLIQ	2.68	MLIQ	2.7	MLIQ	2.72	MLIQ	2.73	MLIQ	2.69
BIG4	2.36	BIG4	2.35	BIG4	2.36	BIG4	2.36	BIG4	2.36
LIQ	2.3	LIQ	2.35	LIQ	2.31	LIQ	2.3	LIQ	2.25
B_DIS	2.24	B_DIS	2.24	B_DIS	2.24	B_DIS	2.24	B_DIS	2.23
MBDIS	1.7	MBDIS	1.78	MBDIS	1.78	MBDIS	1.77	MBDIS	1.64
ROA	1.56	ROA	1.56	ROA	1.51	ROA	1.51	ROA	1.56
ICMW	1.03	ICMW	1.03	ICMW	1.03	ICMW	1.03	ICMW	1.03
Mean VIF	8.41	Mean VIF	8.64	Mean VIF	8.73	Mean VIF	8.73	Mean VIF	8.34

Model (a)-model (d) are from Beatty and Liao (2014), model (e) is from Bushman and Williams (2012)

ARES_a Increase of abnormal loan loss provision model (a) from eight quarters before fraud period to fraud period
ARES_b Increase of abnormal loan loss provision model (b) from eight quarters before fraud period to fraud period
ARES_c Increase of abnormal loan loss provision model (c) from eight quarters before fraud period to fraud period
ARES_d Increase of abnormal loan loss provision model (d) from eight quarters before fraud period to fraud period
ARES_e Increase of abnormal loan loss provision model (e) from eight quarters before fraud period to fraud period
CC Cost of capital: interest expenses divided by total loans
CAR Tier 1 capital divided by weighted average asset risk (%)
B_DIS Bank distress: DeLisle et al. (2007) Z-score
LIQ Liquidity variable: cash and equivalents divided by total assets
MCC Moderating variable model (a/b/c/d/e) abnormal loan loss provision with cost of capital (interest expense divided by total liabilities)
MCAR Moderating variable model (a/b/c/d/e) abnormal loan loss provision with tier 1 capital (%)
MB_DIS Moderating variable model (a/b/c/d/e) abnormal loan loss provision with bank distress (DeLisle et al. (2007) Z-score)
MLIQ Moderating variable model (a/b/c/d/e) abnormal loan loss provision with liquidity (cash & equivalents divided by total assets)
SIZE Natural log of total assetss
BIG4 Dummy variable, 1 if the fiscal year auditor is the Big Four, 0 otherwise
ROA Income divided by average total assets
ICMW Dummy variable, 1 if there is an internal control material weakness, 0 otherwise

Appendix 5 – Current abnormal loan loss provision

Table 20 Current abnormal loan loss provision

Financial statements fraud=$\alpha_0+\alpha_1$ARES(a/b/c/d/e)+ α_2CC+α_3CAR+α_4B_DIS+α_5LIQ+α_6MCC(a/b/c/d/e)+α_7MCAR(a/b/c/d/e)+α_8MB_DIS(a/b/c/d/e)+α_9MLIQ(a/b/c/d/e) +α_{10}BIG4+α_{11}ROA+α_{12}ICMW+error

Variables	Prediction	Model (a) Coefficients (z-statistics)	Model (b) Coefficients (z-statistics)	Model (c) Coefficients (z-statistics)	Model (d) Coefficients (z-statistics)	Model (e) Coefficients (z-statistics)
Hypothesis 1						
ARESa		87.125***				
		(3.65)				
ARESb			93.146***			
			(3.51)			
ARESc				140.126***		
				(3.76)		
ARESd					134.312***	
					(3.59)	
ARESe						125.156***
						(5.79)
Hypotheses 2a						
CC	+	187.795***	187.374***	186.006***	185.797***	189.382***
		(16.32)	(16.30)	(16.20)	(16.19)	(16.43)
CAR	-	-8.658***	-8.827***	-8.598***	-8.601***	-8.525***
		(-5.93)	(-6.05)	(-5.90)	(-5.90)	(-5.84)
B_DIS	-	-0.000**	-0.000**	-0.000**	-0.000**	-0.000*
		(-2.16)	(-2.03)	(-1.97)	(-2.04)	(-1.81)
LIQ	-	4.055***	4.197***	4.250***	4.231***	3.888***
		(5.16)	(5.35)	(5.41)	(5.39)	(4.90)
Control variables						
BIG4	-	0.714***	0.721***	0.734***	0.733***	0.700***
		(10.37)	(10.48)	(10.65)	(10.63)	(10.15)
ROA	-	-29.268***	-29.765***	-31.152***	-31.455***	-22.448***
		(-3.66)	(-3.73)	(-4.00)	(-4.03)	(-2.77)
ICMW	+	0.024	0.013	-0.026	-0.022	-0.002
		(0.10)	(0.06)	(-0.11)	(-0.09)	(-0.01)
constant	?	-2.546***	-2.531***	-2.549***	-2.539***	-2.654***
		(-11.56)	(-11.51)	(-11.60)	(-11.55)	(-12.05)
prob > chi-2		0.0000	0.0000	0.0000	0.0000	0.0000
pseudo r2		0.0945	0.0943	0.0945	0.0944	0.0973
N		9559	9559	9559	9559	9559

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Model (a)-model (d) are from Beatty and Liao (2014), model (e) is from Bushman and Williams (2012)

Definition of the variables:

ARESa	Current abnormal loan loss provision model (a).
ARESb	Current abnormal loan loss provision model (b).
ARESc	Current abnormal loan loss provision model (c).
ARESd	Current abnormal loan loss provision model (d).
ARESe	Current abnormal loan loss provision model (e).
CC	Cost of capital: interest expenses divided by total loans.
CAR	Tier-1 capital divided by weighted average asset risk (%).
B_DIS	Bank distress: DeLisle et al. (2007) Z-score.
LIQ	Liquidity variable: cash and equivalents divided by total assets.
MCC	Moderating variable model (a/b/c/d/e) abnormal loan loss provision with cost of capital (interest expense divided by total liabilities).
MCAR	Moderating variable model (a/b/c/d/e) abnormal loan loss provision with tier 1 capital (%).
MB_DIS	Moderating variable model (a/b/c/d/e) abnormal loan loss provision with bank distress (DeLisle et al. (2007) Z-score).
MLIQ	Moderating variable model (a/b/c/d/e) abnormal loan loss provision with liquidity (cash & equivalents divided by total assets).

BIG4	Dummy variable, 1 if the fiscal year auditor is the Big Four, 0 otherwise.
ROA	Income divided by average total assets.
ICMW	Dummy variable, 1 if there is an internal control material weakness, 0 otherwise.

Table 21 Current abnormal loan loss provision with year fixed effect

Financial statements fraud=$\alpha_0+\alpha_1$ARES(a/b/c/d/e)+ α_2CC+α_3CAR+α_4B_DIS+α_5LIQ+α_6MCC(a/b/c/d/e)+α_7MCAR(a/b/c/d/e)+α_8MB_DIS(a/b/c/d/e)+α_9MLIQ(a/b/c/d/e) +α_{10}BIG4+α_{11}ROA+α_{12}ICMW+error

Variables	Prediction	Model (a) Coefficients (z-statistics)	Model (b) Coefficients (z-statistics)	Model (c) Coefficients (z-statistics)	Model (d) Coefficients (z-statistics)	Model (e) Coefficients (z-statistics)
Hypothesis 1						
ARESa	+	-125.650***				
		(-2.81)				
ARESb	+		-129.877***			
			(-2.68)			
ARESc	+			-63.639		
				(-0.96)		
ARESd	+				-75.207	
					(-1.14)	
ARESe	+					-46.781
						(-1.15)
Hypotheses 2a						
CC	+	63.365	66.382	66.961	66.847	64.418
		(1.45)	(1.51)	(1.53)	(1.52)	(1.47)
CAR	-	-1.449	-1.308	-1.715	-1.735	-1.616
		(-0.53)	(-0.48)	(-0.63)	(-0.64)	(-0.59)
B_DIS	-	0.000	0.000	0.000	0.000	0.000
		(0.27)	(0.16)	(0.19)	(0.18)	(0.25)
LIQ	-	2.983*	2.909*	2.746*	2.755*	2.840*
		(1.90)	(1.86)	(1.76)	(1.77)	(1.81)
Control variables						
BIG4	-	0.380	0.364	0.385	0.388	0.378
		(1.58)	(1.50)	(1.59)	(1.60)	(1.56)
ROA	-	-18.914*	-18.741*	-14.177	-14.463	-15.162
		(-1.70)	(-1.69)	(-1.28)	(-1.31)	(-1.36)
ICMW	+	0.535	0.543	0.480	0.490	0.487
		(1.61)	(1.63)	(1.44)	(1.47)	(1.46)
Year fixed effect						
Fiscal Year=2003		0.000	0.000	0.000	0.000	0.000
		(.)	(.)	(.)	(.)	(.)
Fiscal Year=2004		0.048	0.082	0.101	0.101	0.096
		(0.11)	(0.19)	(0.23)	(0.23)	(0.22)
Fiscal Year=2005		-0.385	-0.339	-0.346	-0.347	-0.340
		(-0.88)	(-0.77)	(-0.79)	(-0.79)	(-0.78)
Fiscal Year=2006		1.199***	1.230***	1.232***	1.230***	1.264***
		(2.91)	(2.98)	(2.99)	(2.99)	(3.06)
Fiscal Year=2007		2.228***	2.251***	2.224***	2.225***	2.265***
		(5.29)	(5.34)	(5.28)	(5.28)	(5.35)
Fiscal Year=2008		3.571***	3.589***	3.489***	3.496***	3.526***
		(9.21)	(9.22)	(9.01)	(9.03)	(9.02)
Fiscal Year=2009		2.438***	2.432***	2.255***	2.269***	2.310***
		(6.29)	(6.26)	(5.88)	(5.91)	(5.89)
Fiscal Year=2010		1.137***	1.115***	1.026**	1.032**	1.086***
		(2.82)	(2.77)	(2.56)	(2.57)	(2.67)
Fiscal Year=2011		-0.190	-0.122	-0.180	-0.177	-0.183
		(-0.42)	(-0.27)	(-0.39)	(-0.39)	(-0.40)
Fiscal Year=2012		0.552	0.612	0.544	0.546	0.540
		(1.28)	(1.42)	(1.26)	(1.27)	(1.25)
Fiscal Year=2013		-0.728	-0.677	-0.739	-0.737	-0.727

	(-1.41)	(-1.31)	(-1.43)	(-1.43)	(-1.41)
Fiscal Year=2014	0.705	0.753*	0.730*	0.729*	0.748*
	(1.63)	(1.75)	(1.69)	(1.69)	(1.73)
Fiscal Year=2015	1.100***	1.139***	1.124***	1.123***	1.164***
	(2.61)	(2.71)	(2.67)	(2.67)	(2.76)
Prob > chi2	0.0000	0.0000	0.0000	0.0000	0.0000
LR chi2(20)	1138.07	1137.27	1130.86	1131.23	1131.27
N	6840	6840	6840	6840	6840

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Model (a)-model (d) are from Beatty and Liao (2014), model (e) is from Bushman and Williams (2012)

Definition of the variables:

ARES_a	Increase of abnormal loan loss provision model (a) from eight quarters before fraud period to fraud period.
ARES_b	Increase of abnormal loan loss provision model (b) from eight quarters before fraud period to fraud period.
ARES_c	Increase of abnormal loan loss provision model (c) from eight quarters before fraud period to fraud period.
ARES_d	Increase of abnormal loan loss provision model (d) from eight quarters before fraud period to fraud period.
ARES_e	Increase of abnormal loan loss provision model (e) from eight quarters before fraud period to fraud period.
CC	Cost of capital: interest expenses divided by total loans.
CAR	Tier-1 capital divided by weighted average asset risk (%).
B_DIS	Bank distress: DeLisle et al. (2007) Z-score.
LIQ	Liquidity variable: cash and equivalents divided by total assets.
MCC	Moderating variable model (a/b/c/d/e) abnormal loan loss provision with cost of capital (interest expense divided by total liabilities).
MCAR	Moderating variable model (a/b/c/d/e) abnormal loan loss provision with tier 1 capital (%).
MB_DIS	Moderating variable model (a/b/c/d/e) abnormal loan loss provision with bank distress (DeLisle et al. (2007) Z-score).
MLIQ	Moderating variable model (a/b/c/d/e) abnormal loan loss provision with liquidity (cash & equivalents divided by total assets).
BIG4	Dummy variable, 1 if the fiscal year auditor is the Big Four, 0 otherwise.
ROA	Income divided by average total assets.
ICMW	Dummy variable, 1 if there is an internal control material weakness, 0 otherwise.

Appendix 6 – Bushman and Williams (2012)

Table 22 OLS regression Bushman and Williams (2012)

	Prediction	Bushman and Williams (2012) model (e)	Bushman and Williams (2012) model (f)
		b/t	b/t
ΔNPA t+1	+	0.013***	0.013***
		(2.67)	(2.64)
ΔNPA t	+	0.044***	0.043***
		(5.32)	(5.27)
ΔNPA t-1	+	0.044***	0.044***
		(9.14)	(8.91)
ΔNPA t-2	+	0.048***	0.048***
		(9.12)	(8.86)
SIZE_1	+	0.000***	0.000***
		(3.87)	(3.98)
CH_GDP	+	-0.038***	-0.036***
		(-15.78)	(-15.10)
EBP	-		-0.026**
			(-2.04)
CAR	-		-0.002
			(-0.90)
constant		0.002***	0.002***
		(4.78)	(3.74)
R-sqr		0.232	0.238
N		9715	9715

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Standard errors are clustered at the bank level. Model (a) is used in the main regression analysis. Model (b) is the original model of Bushman and Williams (2012).

Model (e): $LLP_{j,t} = \alpha_0 + \alpha_1 \Delta NPA_{t+1} + \alpha_2 \Delta NPA_t + \alpha_3 \Delta NPA_{t-1} + \alpha_4 \Delta NPA_{t-2} + \alpha_5 \Delta SIZE_{t-1} + \alpha_6 \Delta GDP_t + \varepsilon_{j,t}$

Model (f): $LLP_{j,t} = \alpha_0 + \alpha_1 \Delta NPA_{t+1} + \alpha_2 \Delta NPA_t + \alpha_3 \Delta NPA_{t-1} + \alpha_4 \Delta NPA_{t-2} + \alpha_5 \Delta SIZE_{t-1} + \alpha_6 \Delta GDP_t + \alpha_7 \mathrm{EBP}_t + \alpha_8 CAR_t + \varepsilon_{j,t}$

Definition of the variables:

LLP the loan loss provision divided by lagged total loans.
ΔNPA the change of non-performing assets divided by lagged total loans.
SIZE the natural log of total assets.
ΔLoan the change of total loan divided by lagged total loans.
ΔGDP the change of Gross Domestic Product over the quarter.
EBP Earnings before loan loss provision.
CAR Tier 1 capital divided by weighted average asset risk (%)

Table 23 Logit regression with Bushman and Williams (2012) model

$$\text{Financial statements fraud}=\alpha_0+\alpha_1\text{ARES\_(e/f)}+\alpha_2\text{CC}+\alpha_3\text{CAR}+\alpha_4\text{B\_DIS}+\alpha_5\text{LIQ}+\alpha_6\text{MCC(e/f)}+\alpha_7\text{MCAR(e/f)}+\alpha_8\text{MB\_DIS(e/f)}+\alpha_9\text{MLIQ(e/f)}+\alpha_{10}\text{BIG4}+\alpha_{11}\text{ROA}+\alpha_{12}\text{ICMW}+\text{error}$$

	Prediction	Model (e) b/z	Model (f) b/z	Model (e) b/z	Model (f) b/z
Hypothesis 1					
ARES_e	+	181.813***		-42.611	
		(5.49)		(-1.08)	
ARES_f	+		123.049		-26.652
			(1.30)		(-0.21)
Hypotheses 2a					
CC	+	216.031***	180.212***	99.915**	70.203
		(14.19)	(15.30)	(2.09)	(1.59)
CAR	-	-10.003***	-8.759***	-5.424	-2.037
		(-5.52)	(-5.72)	(-1.64)	(-0.74)
B_DIS	-	-0.002***	-0.001***	-0.001	-0.000
		(-7.29)	(-3.77)	(-1.32)	(-0.02)
LIQ	-	6.995***	4.502***	5.825**	2.867*
		(5.59)	(5.70)	(2.52)	(1.78)
Hypotheses 2b					
CCe	+	-24984.378***	-7082.763	-26874.421**	-14356.690
		(-3.42)	(-1.04)	(-2.24)	(-1.62)
CAR_e	+	407.200	-11.198	1244.401*	351.796
		(0.94)	(-0.02)	(1.78)	(0.44)
BDISe	+	1.543***	0.819***	0.568*	0.313*
		(7.83)	(5.36)	(1.90)	(1.80)
LIQe	+	-1403.964***	-582.458*	-1323.486*	-139.090
		(-2.91)	(-1.70)	(-1.76)	(-0.28)
Control variables					
BIG4	-	0.681***	0.697***	0.398	0.366
		(9.76)	(10.07)	(1.64)	(1.51)
ROA	-	-31.827***	-29.329***	-21.827*	-17.052
		(-3.91)	(-3.71)	(-1.95)	(-1.54)
ICMW	+	-0.084	-0.062	0.509	0.471
		(-0.36)	(-0.26)	(1.54)	(1.42)
				(-0.79)	(-0.77)
Year fixed effect					
Fiscal Year=2006				1.233***	1.224***
				(2.96)	(2.96)
Fiscal Year=2007				2.237***	2.234***
				(5.25)	(5.26)
Fiscal Year=2008				3.482***	3.486***
				(8.78)	(8.89)
Fiscal Year=2009				2.253***	2.253***
				(5.70)	(5.83)
Fiscal Year=2010				0.999**	1.021**
				(2.44)	(2.53)
Fiscal Year=2011				-0.283	-0.203
				(-0.61)	(-0.44)
Fiscal Year=2012				0.495	0.561
				(1.13)	(1.29)
Fiscal Year=2013				-0.772	-0.766
				(-1.48)	(-1.47)
Fiscal Year=2014				0.756*	0.741*
				(1.72)	(1.70)
Fiscal Year=2015				1.036**	1.123***
				(2.39)	(2.64)
constant	?	-2.310***	-2.352***		
		(-10.39)	(-10.54)		
R-sqr		0.1154	0.1068	-	-
N		9551	9551	6839	6839

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Model (e) is from Bushman and Williams (2012) model used in the main regression. Model (f) is Bushman and Williams original model.

Definition of the variables:

ARES_e	Increase of abnormal loan loss provision model (e) from eight quarters before fraud period to fraud period
ARES_f	Increase of abnormal loan loss provision model (f) from eight quarters before fraud period to fraud period
CC	Cost of capital: interest expenses divided by total loans

CAR	Tier 1 capital divided by weighted average asset risk (%)
B_DIS	Bank distress: DeLisle et al. (2007) Z-score.
LIQ	Liquidity variable: cash and equivalents divided by total assets
MCC	Moderating variable model (a/b/c/d/e) abnormal loan loss provision with cost of capital (interest expense divided by total liabilities)
MCAR	Moderating variable model (a/b/c/d/e) abnormal loan loss provision with tier 1 capital (%)
MB_DIS	Moderating variable model (a/b/c/d/e) abnormal loan loss provision with bank distress (DeLisle et al. (2007) Z-score)
MLIQ	Moderating variable model (a/b/c/d/e) abnormal loan loss provision with liquidity (cash & equivalents divided by total assets)
BIG4	Dummy variable, 1 if the fiscal year auditor is the Big Four, 0 otherwise
ROA	Income divided by average total assets
ICMW	Dummy variable, 1 if there is an internal control material weakness, 0 otherwise

Appendix 7 – Bank distress

Table 24 List of Z-scores

No.	Literature	Z-score
1.	Maecheler et al. (2007) use in the main analysis	$Bank\ distress_{j,t} = \dfrac{3\ years\ moving\ average \dfrac{Equity_{j,t}}{Total\ assets_{j,t}} + 3\ years\ moving\ average\ ROA_{j,t}}{3\ years\ \sigma\ ROA_{j,t}}$
2.	Boyd and Graham (1986)	$Bank\ distress_{j,t} = \dfrac{\dfrac{Equity_{j,t}}{Total\ assets_{j,t}} + ROA_{j,t}}{3\ years\ \sigma\ ROA_{j,t}}$
3.	Boyd et al. (2007)	$Bank\ distress_{j,t} = \dfrac{3\ years\ moving\ average \dfrac{Equity_{j,t}}{Total\ assets_{j,t}} + current\ ROA_{j,t}}{3\ years\ \sigma\ ROA_{j,t}}$

ROA Earnings divided by average total assets.
Current ROA Earnings divided by current total assets.

Table 25 Logit regression with additional Z-scores

Financial statements fraud=$\alpha_0+\alpha_1$ARES_a+ α_2CC+α_3CAR+α_4B_DIS(main model/2/3)+α_5LIQ+α_6MCC+α_7MCAR +α_8MB_DIS(main model/2/3)+α_9MLIQ+α_{10}BIG4+α_{11}ROA+α_{12}ICMW+error

		Random effect			Year fixed effect		
	Prediction	Bank distress main model b/z	Bank distress model 2 b/z	Bank distress model 3 b/z	Bank distress main model b/z	Bank distress model 2 b/z	Bank distress model 3 b/z
ARES_a	+	132.549***	133.216***	132.427***	-96.416***	-96.495***	-96.422***
		(6.67)	(6.70)	(6.67)	(-3.28)	(-3.28)	(-3.28)
CC	+	182.074***	181.648***	182.036***	70.931	70.896	70.938
		(15.75)	(15.74)	(15.74)	(1.62)	(1.62)	(1.62)
T1CAP	-	-8.077***	-8.048***	-8.078***	-1.746	-1.753	-1.745
		(-5.49)	(-5.47)	(-5.49)	(-0.64)	(-0.64)	(-0.64)
B_DIS(main2/3)	-	-0.000**	-0.000**	-0.000**	0.000	0.000	0.000
		(-2.28)	(-2.04)	(-2.31)	(0.31)	(0.40)	(0.32)
LIQ	-	4.204***	4.196***	4.203***	2.768*	2.772*	2.768*
		(5.36)	(5.35)	(5.36)	(1.76)	(1.77)	(1.76)
BIG4	-	0.720***	0.719***	0.720***	0.368	0.369	0.368
		(10.44)	(10.42)	(10.44)	(1.52)	(1.53)	(1.52)
ROA	-	-22.888***	-23.285***	-22.792***	-17.877	-18.010	-17.897
		(-2.89)	(-2.94)	(-2.88)	(-1.62)	(-1.63)	(-1.62)
ICMW	+	-0.021	-0.017	-0.022	0.551*	0.551*	0.551*
		(-0.09)	(-0.07)	(-0.09)	(1.65)	(1.65)	(1.65)
year fixed effect							
Fiscal Year=2003					0.000	0.000	0.000
					(.)	(.)	(.)
Fiscal Year=2004					0.078	0.077	0.078
					(0.18)	(0.18)	(0.18)
Fiscal Year=2005					-0.367	-0.371	-0.368
					(-0.84)	(-0.84)	(-0.84)

Fiscal Year=2006					1.198***	1.194***	1.197***
					(2.90)	(2.89)	(2.90)
Fiscal Year=2007					2.241***	2.238***	2.241***
					(5.31)	(5.30)	(5.31)
Fiscal Year=2008					3.583***	3.582***	3.583***
					(9.22)	(9.22)	(9.22)
Fiscal Year=2009					2.429***	2.430***	2.429***
					(6.29)	(6.30)	(6.30)
Fiscal Year=2010					1.058***	1.059***	1.058***
					(2.64)	(2.64)	(2.64)
Fiscal Year=2011					-0.375	-0.374	-0.375
					(-0.81)	(-0.81)	(-0.81)
Fiscal Year=2012					0.458	0.458	0.458
					(1.06)	(1.06)	(1.06)
Fiscal Year=2013					-0.721	-0.722	-0.721
					(-1.40)	(-1.40)	(-1.40)
Fiscal Year=2014					0.715*	0.712*	0.715*
					(1.65)	(1.65)	(1.65)
Fiscal Year=2015					1.116***	1.113***	1.116***
					(2.64)	(2.64)	(2.64)
constant		-2.499***	-2.509***	-2.497***			
		(-11.58)	(-11.67)	(-11.58)			
R-sqr							
N		9551	9553	9551	6839	6839	6839

Main model uses DeLisle et al. (2007) Z-score. Model 2 used Boyd and Graham et al. (1986) Z-score. Model 3 used Boyd et al. (2007) Z-score.

Definition of the variables:

ARES_a	Increase of abnormal loan loss provision model (a) from eight quarters before fraud period to fraud period.
CC	Cost of capital: interest expenses divided by total loans.
CAR	Tier 1 capital divided by weighted average asset risk (%).
B_DIS	Bank distress
LIQ	Liquidity variable: cash and equivalents divided by total assets.
BIG4	Dummy variable, 1 if the fiscal year auditor is the Big Four, 0 otherwise.
ROA	Income divided by average total assets.
ICMW	Dummy variable, 1 if there is an internal control material weakness, 0 otherwise.

Table 26 Logit regression with additional Z-scores with year fixed effect

Financial statements fraud=$\alpha_0+\alpha_1$ARES_a+ α_2CC+α_3CAR+α_4B_DIS(main model/2/3)+α_5LIQ+α_6MCC+α_7MCAR +α_8MB_DIS(main model/2/3)+α_9MLIQ+α_{10}BIG4+α_{11}ROA+α_{12}ICMW+error

		Random effect			Year fixed effect		
	Prediction	Bank distress main model b/z	Bank distress model 2 b/z	Bank distress model 3 b/z	Bank distress main model b/z	Bank distress model 2 b/z	Bank distress model 3 b/z
ARES_a	+	-19.880	-8.628	-19.742	-277.193*	-272.367*	-277.051*
		(-0.17)	(-0.08)	(-0.17)	(-1.79)	(-1.76)	(-1.79)
CC	+	180.229***	179.776***	180.189***	67.788	68.526	67.731
		(15.44)	(15.43)	(15.44)	(1.54)	(1.56)	(1.54)
T1CAP	-	-8.951***	-8.934***	-8.948***	-2.201	-2.200	-2.199
		(-5.86)	(-5.85)	(-5.86)	(-0.79)	(-0.79)	(-0.79)
B_DIS(main2/3)	-	-0.000**	-0.000**	-0.000**	0.000	0.000	0.000
		(-2.33)	(-2.07)	(-2.36)	(0.34)	(0.39)	(0.35)
LIQ	-	4.647***	4.631***	4.645***	2.870*	2.862*	2.869*
		(5.87)	(5.86)	(5.87)	(1.79)	(1.79)	(1.79)
MCC	+	3786.704	3936.729	3889.225	-1933.649	-1173.776	-1921.426

		(0.49)	(0.51)	(0.50)	(-0.17)	(-0.11)	(-0.17)
MCARa	+	704.698	631.360	703.376	1274.175	1250.569	1271.428
		(0.97)	(0.87)	(0.96)	(1.27)	(1.25)	(1.27)
MBDIS2(main2/3)	+	0.772***	0.730***	0.770***	0.365*	0.295	0.367*
		(4.91)	(4.76)	(4.90)	(1.86)	(1.55)	(1.88)
MLIQ	+	-391.419	-394.492	-393.185	39.969	42.953	39.488
		(-0.90)	(-0.91)	(-0.90)	(0.07)	(0.07)	(0.06)
BIG4	-	0.734***	0.732***	0.734***	0.371	0.371	0.371
		(10.60)	(10.56)	(10.60)	(1.53)	(1.53)	(1.53)
ROA	-	-31.468***	-31.827***	-31.519***	-23.401**	-22.882**	-23.526**
		(-3.91)	(-3.95)	(-3.92)	(-2.07)	(-2.02)	(-2.08)
ICMW	+	-0.061	-0.059	-0.061	0.564*	0.563*	0.564*
		(-0.26)	(-0.25)	(-0.26)	(1.70)	(1.69)	(1.70)
Fiscal Year=2003					0.000	0.000	0.000
					(.)	(.)	(.)
Fiscal Year=2004					0.092	0.086	0.092
					(0.21)	(0.20)	(0.21)
Fiscal Year=2005					-0.343	-0.351	-0.343
					(-0.78)	(-0.80)	(-0.78)
Fiscal Year=2006					1.241***	1.229***	1.241***
					(2.99)	(2.96)	(2.99)
Fiscal Year=2007					2.257***	2.252***	2.257***
					(5.34)	(5.33)	(5.34)
Fiscal Year=2008					3.585***	3.587***	3.585***
					(9.18)	(9.19)	(9.18)
Fiscal Year=2009					2.438***	2.437***	2.438***
					(6.31)	(6.31)	(6.31)
Fiscal Year=2010					1.068***	1.070***	1.068***
					(2.66)	(2.66)	(2.65)
Fiscal Year=2011					-0.364	-0.360	-0.364
					(-0.78)	(-0.77)	(-0.78)
Fiscal Year=2012					0.464	0.468	0.464
					(1.06)	(1.07)	(1.06)
Fiscal Year=2013					-0.731	-0.725	-0.731
					(-1.41)	(-1.40)	(-1.41)
Fiscal Year=2014					0.756*	0.750*	0.756*
					(1.73)	(1.72)	(1.73)
Fiscal Year=2015					1.150***	1.144***	1.150***
					(2.70)	(2.69)	(2.70)
constant		-2.384***	-2.393***	-2.383***			
		(-10.75)	(-10.82)	(-10.75)			
R-sqr							
N		9551	9553	9551	6839	6839	6839

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Main model uses DeLisle et al. (2007) Z-score. Model 2 used Boyd and Graham et al. (1986) Z-score. Model 3 used Boyd et al. (2007) Z-score model.

Definition of the variables:

ARES_a	Increase of abnormal loan loss provision model (a) from eight quarters before fraud period to fraud period.
CC	Cost of capital: interest expenses divided by total loans.
CAR	Tier 1 capital divided by weighted average asset risk (%).
B_DIS	Bank distress.
LIQ	Liquidity variable: cash and equivalents divided by total assets.
MCC	Moderating variable model (a/b/c/d/e) abnormal loan loss provision with cost of capital (interest expense divided by total liabilities).
MCAR	Moderating variable model (a/b/c/d/e) abnormal loan loss provision with tier 1 capital (%).
MB_DIS	Moderating variable model (a/b/c/d/e) abnormal loan loss provision with bank distress.
MLIQ	Moderating variable model (a/b/c/d/e) abnormal loan loss provision with liquidity (cash & equivalents divided by total assets).
BIG4	Dummy variable, 1 if the fiscal year auditor is the Big Four, 0 otherwise.
ROA	Income divided by average total assets.
ICMW	Dummy variable, 1 if there is an internal control material weakness, 0 otherwise.

Appendix 8 – Net stable funding ratio

Equation 14
Net stable funding ratio (NSFR)

$$NSFR = \frac{Available\ stable\ fund\ (ASF)}{Required\ stable\ funding\ (RSF)}$$

$$ASF = Equity + Total\ LT\ Funding + \binom{Customer\ deposits}{savings\ Term} * 0.9 + \binom{Customer\ deposits}{current} * 0.9 + \binom{Other\ depostis\ and}{ST\ borrowing} * 0.5$$

$$RSF = Other\ assets + (Goverment\ securities + OBS\ items) * 0.05 + (Other\ securities + Loans\ advances\ to\ banks) * 0.5 + (Mortgage\ loans) * 0.65 + (Retail\ and\ corporate\ loans) * 0.85$$

Table 27 Logit regression with NSFR

Financial statements fraud=α_0+α_1ARES_a+ α_2CC+α_3CAR+α_4B_DIS(main model/2/3)+α_5LIQ+α_6MCC+α_7MCAR +α_8MB_DIS(main model/2/3)+α_9MLIQ+α_{10}BIG4+α_{11}ROA+α_{12}ICMW+error

	Prediction	random effect		year fixed effect	
		liquidity ratio	NSFR	liquidity ratio	NSFR
		b/z	b/z	b/z	b/z
hypothesis 1					
ARES_a	+	121.653***	123.464***	-107.218***	-106.835***
		(5.62)	(5.70)	(-3.23)	(-3.21)
hypotheses 2a					
CC	+	181.810***	159.596***	57.413	59.664
		(14.63)	(13.05)	(1.15)	(1.20)
T1CAP	-	-7.739***	-8.146***	-2.066	-2.115
		(-4.83)	(-5.03)	(-0.68)	(-0.70)
B_DIS(main2/3)	-	-0.000	-0.000	0.000	0.000
		(-1.22)	(-1.30)	(0.97)	(0.93)
LIQ	-	4.072***		1.973	
		(4.99)		(1.09)	
NSFR	-		0.198***		0.353*
			(4.49)		(1.95)
control variables					
BIG4	-	0.783***	0.686***	0.468	0.481*
		(10.08)	(8.52)	(1.63)	(1.67)
ROA	-	-26.409***	-29.843***	-17.968	-19.263
		(-3.00)	(-3.39)	(-1.43)	(-1.53)
ICMW	+	-0.010	0.013	0.405	0.399
		(-0.04)	(0.05)	(1.16)	(1.13)
fixed year effect					
Fiscal Year=2003				0.000	0.000
				(.)	(.)
Fiscal Year=2004				-0.529	-0.641
				(-0.90)	(-1.07)
Fiscal Year=2005				-0.598	-0.604
				(-1.13)	(-1.14)

Fiscal Year=2006				0.799	0.800
				(1.56)	(1.57)
Fiscal Year=2007				1.807***	1.789***
				(3.47)	(3.45)
Fiscal Year=2008				3.007***	2.983***
				(6.27)	(6.23)
Fiscal Year=2009				2.092***	2.114***
				(4.43)	(4.47)
Fiscal Year=2010				0.573	0.620
				(1.18)	(1.27)
Fiscal Year=2011				-0.848	-0.787
				(-1.58)	(-1.46)
Fiscal Year=2012				-0.212	-0.131
				(-0.41)	(-0.25)
Fiscal Year=2013				-1.296**	-1.205**
				(-2.21)	(-2.05)
Fiscal Year=2014				0.114	0.213
				(0.22)	(0.41)
Fiscal Year=2015				0.481	0.577
				(0.95)	(1.13)
constant		-2.572***	-2.349***		
		(-10.77)	(-9.95)		
R-sqr					
N		7856	7856	5510	5510

Table 28 Logit regression NSFR with year fixed effect

Financial statements fraud=$\alpha_0+\alpha_1$ARES_a+ α_2CC+α_3CAR+α_4B_DIS(main model/2/3)+α_5LIQ+α_6MCC+α_7MCAR +α_8MB_DIS(main model/2/3)+α_9MLIQ+α_{10}BIG4+α_{11}ROA+α_{12}ICMW+error

		random effect		year fixed effect	
	Prediction	liquidity ratio	NSFR	liquidity ratio	NSFR
		b/z	b/z	b/z	b/z
hypothesis 1					
ARES_a	+	-59.562	-74.530	-301.442*	-212.555
		(-0.47)	(-0.64)	(-1.81)	(-1.26)
hypotheses 2a					
CC	+	179.573***	155.898***	50.295	45.065
		(14.29)	(12.58)	(1.00)	(0.90)
T1CAP	-	-8.930***	-9.200***	-2.730	-2.296
		(-5.35)	(-5.46)	(-0.89)	(-0.74)
B_DIS(main2/3)	-	-0.000	-0.000	0.000	0.000
		(-1.41)	(-1.50)	(1.01)	(0.92)
LIQ	-	4.557***		2.278	
		(5.53)		(1.24)	
NSFR	-		0.200***		0.294
			(4.47)		(1.55)
hypotheses 2b					
MCC	+	-505.592	1658.597	-6086.985	-5329.161
		(-0.06)	(0.21)	(-0.50)	(-0.44)
MCARa	+	1105.313	1143.143	1536.037	1412.822
		(1.38)	(1.44)	(1.43)	(1.30)
MBDISa	+	0.893***	0.883***	0.537**	0.542**
		(5.12)	(5.06)	(2.44)	(2.45)
MLIQa	+	-427.993		-331.925	
		(-0.92)		(-0.51)	
MNSFRa	+		-21.904		-83.141**
			(-0.73)		(-2.20)
control variables					
BIG4	-	0.734***	0.732***	0.371	0.371

		(10.60)	(10.56)	(1.53)	(1.53)
ROA	-	-31.468***	-31.827***	-23.401**	-22.882**
		(-3.91)	(-3.95)	(-2.07)	(-2.02)
ICMW	+	-0.061	-0.059	0.564*	0.563*
		(-0.26)	(-0.25)	(1.70)	(1.69)
Fiscal Year=2003				0.000	0.000
				(.)	(.)
Fiscal Year=2004				-0.543	-0.661
				(-0.92)	(-1.10)
Fiscal Year=2005				-0.557	-0.557
				(-1.04)	(-1.04)
Fiscal Year=2006				0.855*	0.878*
				(1.66)	(1.70)
Fiscal Year=2007				1.837***	1.849***
				(3.51)	(3.53)
Fiscal Year=2008				3.000***	2.975***
				(6.21)	(6.14)
Fiscal Year=2009				2.099***	2.087***
				(4.42)	(4.37)
Fiscal Year=2010				0.576	0.591
				(1.18)	(1.20)
Fiscal Year=2011				-0.841	-0.876
				(-1.54)	(-1.58)
Fiscal Year=2012				-0.215	-0.175
				(-0.41)	(-0.33)
Fiscal Year=2013				-1.330**	-1.264**
				(-2.25)	(-2.13)
Fiscal Year=2014				0.152	0.218
				(0.29)	(0.42)
Fiscal Year=2015				0.516	0.569
				(1.00)	(1.10)
constant		-2.408***	-2.181***		
		(-9.81)	(-8.95)		
R-sqr					
N		7856	7856	5510	5510

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Definition of the variables:

ARES_a	Increase of abnormal loan loss provision model (a) from eight quarters before fraud period to fraud period
ARES_f	Increase of abnormal loan loss provision model (f) from eight quarters before fraud period to fraud period
CC	Cost of capital: interest expenses divided by total loans
CAR	Tier 1 capital divided by weighted average asset risk (%)
B_DIS	Bank distress: DeLisle et al. (2007) Z-score.
LIQ	Liquidity variable: cash and equivalents divided by total assets
NSFR	Net stable funding ratio
MCC	Moderating variable model (a/b/c/d/e) abnormal loan loss provision with cost of capital (interest expense divided by total liabilities)
MCAR	Moderating variable model (a/b/c/d/e) abnormal loan loss provision with tier 1 capital (%)
MB_DIS	Moderating variable model (a/b/c/d/e) abnormal loan loss provision with bank distress (DeLisle et al. (2007) Z-score)
MLIQ	Moderating variable model (a/b/c/d/e) abnormal loan loss provision with liquidity (cash & equivalents divided by total assets)
BIG4	Dummy variable, 1 if the fiscal year auditor is the Big Four, 0 otherwise
ROA	Income divided by average total assets
ICMW	Dummy variable, 1 if there is an internal control material weakness, 0 otherwise

Appendix 9 – F-score prediction results approach

Equation 15
Dechow et al. (2011) F-score results approach

Unconditional probability =
$\alpha_0+\alpha_1 *$ mean ARES(a/b/c/d/e)+ $\alpha_{2}*$ mean CC+α_3 * mean CAR+$\alpha_4 *$ mean B_DIS+
$\alpha_5 *$ mean LIQ+$\alpha_6$meanROA

$$Probability = \frac{e^{Predicted\ Value}}{(1 + e^{(Predicted\ Value)})}$$

$$F - score = \frac{Probability}{Unconditional\ probability}$$