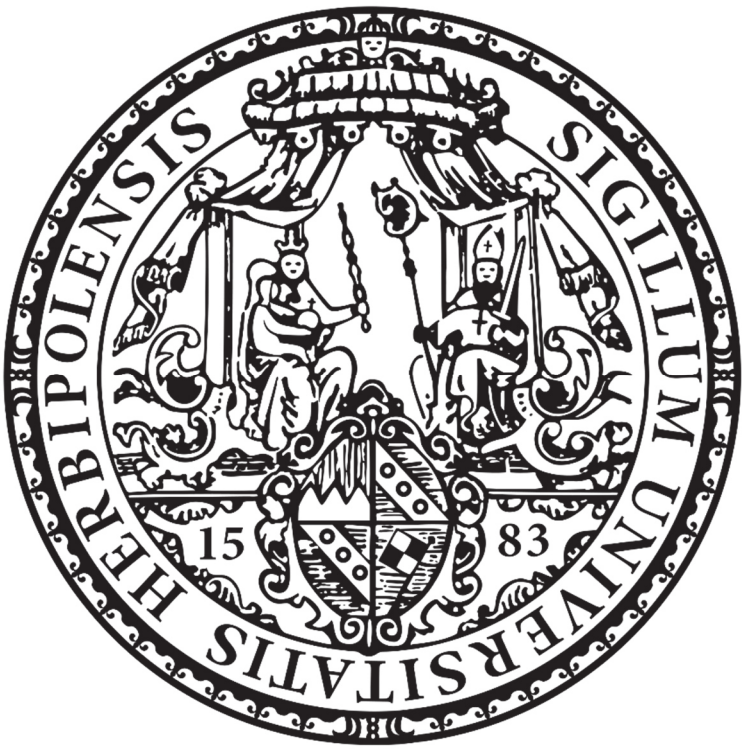
Inauguraldissertation
für die Verleihung des akademischen Grades eines Doktors
der Wirtschaftswissenschaften
der Julius-Maximilians-Universität

Earnings Management in the Context of Earnings Quality, Corporate Governance and Corporate Social Responsibility



Daniel Norbert Schaupp
Würzburg

First supervisor: Prof. Dr. Andrea Szczesny

Second supervisor: Prof. Dr. Hansrudi Lenz

Acknowledgments

Initially, I would like to express profound gratitude and appreciation to my first supervisor, Professor Dr. Andrea Szczesny, and likewise to my second supervisor, Professor Dr. Hansrudi Lenz, who co-authored one paper, which is one part of this doctoral thesis. Professor Dr. Szczesny has been a great mentor and I am very grateful for the opportunity to write my doctoral thesis at her chair. What is more, every discussion with both Professor Dr. Szczesny and Professor Dr. Lenz has been fruitful and majorly developed my thinking. In addition, I am very thankful for the support by Professor Dr. Szczesny in gaining diverse experiences, e.g., the participation in several conferences and workshops around Europe and the world and the opportunity to stay at the University of Melbourne as a visiting Ph.D. student. I am also very grateful to my second supervisor, Professor Dr. Lenz, for the great collaboration and valuable advice and suggestions on how to develop and refine my research.

Furthermore, I would like to thank Dr. Markus Stralla, for the collaboration on one paper, which is part of this doctoral thesis, as well as the intensive discussions and debates on our ideas and understanding of research. Concurrent striving for excellence in what we do has favoured my development and I greatly appreciate him as a colleague and friend. Likewise, I am very grateful for the possibility to work together with Dr. Sven Hörner, Dr. Johannes Reeg, and Dr. Oliver Unger. You were excellent colleagues and the way we worked together as a team was exceptional.

What is important, I want to thank my entire family, namely Monika Schaupp, Klaus Schaupp, Julian Schaupp, for the persistent trust and support in every circumstance. Knowing you behind me has encouraged every undertaking, while you persistently reminded me of what is important in life. You inspire me in what I do, and my gratitude cannot be expressed in words. To conclude, I would like to thank one special human being, Ina Deisböck. You persistently strengthened me and encouraged me in pursuing my work. Through all the good and particularly the tough times we had to face, I found a deepest connection with you and I am sure that together we are able to overcome everything. I cannot express in words what you mean to me.

Zusammenfassung

Steigende Geschwindigkeit und Komplexität von Geschäftsprozessen und –beziehungen verändern fundamental die Ansprüche an umfassende und werthaltige finanzielle und nicht-finanzielle Informationen über Unternehmen. Relevante Informationen in einem Geschäftsbericht müssen daher zunehmend Vorhersagecharakter besitzen, um ihre Daseinsberechtigung nicht zu verlieren. Die adäquate Lösung existenter Prinzipal-Agenten-Konflikte bedarf zugleich einer detaillierten und objektiven Einschätzung der Leistung des Agenten, was zusätzlich zur Wertrelevanz ein Mindestmaß an Verlässlichkeit der berichteten Informationen bedingt. In diesem Spannungsfeld kommt den Standardsettern die Aufgabe zu, Rahmenbedingungen zu schaffen, die es dem Management ermöglichen, Spielräume bei der Erstellung der Berichte zu nutzen, um diesen Qualitätsansprüchen an Informationen über die aktuelle und zukünftige Lage des Unternehmens gerecht zu werden. Diese bewusst gesetzten Möglichkeiten der Einflussnahme durch das Management begreifen jedoch auch signifikante Risiken der opportunistisch motivierten Beeinflussung der Adressaten, welche unter der Überschrift Bilanzpolitik subsumiert werden können. Die in der empirischen Forschung gängige Klassifizierung unterscheidet zwischen der buchmäßigen und realen Bilanzpolitik, wobei letztere im Vergleich eine gezielte Beeinflussung von Geschäftsvorfällen und nicht deren ex-post Abbildung begreift.

Kapitel 2 der vorliegenden Dissertation propagiert eine umfassende Auseinandersetzung mit der empirischen Messung von Bilanzpolitik. Das Fehlen einer zufriedenstellenden Ermittlung des aggregierten Ausmaßes von Bilanzpolitik erfordert hierbei die Anwendung verschiedener Regressionsmodelle. Im speziellen Kontext der empirischen Analyse von Bilanzpolitik im Bankensektor besteht diesbezüglich eine ausgeprägte Heterogenität der verwendeten Modelle. Zugleich existiert nach wie vor keine systematische Analyse der Ansätze und Modellierungsarten von *Loan Loss Provisions*, welche als elementares Vehikel der diskretionären Einflussnahme im Fokus der einschlägigen Studien steht. Daher untersucht der Verfasser in Kapitel 2.1 prävalente Modellierungsansätze, erstellt einen Baukasten verschiedener Regressionsparameter und testet eine Vielzahl an Varianten im Hinblick auf die Validität der ermittelten bilanzpolitischen Größen. Eine derartige

Analyse ist von elementarer Wichtigkeit für die Validität der Ergebnisse und der gezogenen Schlüsse aus empirischen Studien im Bankensektor und liefert daher einen wichtigen Beitrag zur Entwicklung und Einordnung vergangener und zukünftiger Analysen. Die Ergebnisse der zur Anwendung gebrachten Analysen und Tests zeigen, dass eine Vielzahl der existierenden Modelle eine ausreichende Passgenauigkeit aufweist. Jedoch sind gewisse Regressoren weniger wichtig als zunächst angenommen. So können die Regressoren der *Non-Performing Loans* die Modellierung stark verbessern, während *Loan Loss Reserves* und *Net Charge-Offs* nur einen kleinen Beitrag leisten können. Im weiteren Verlauf lässt sich zudem eine ausgeprägte Nichtlinearität einzelner Regressoren aufzeigen, sowie dass ein Wechsel auf ein Modell mit Berücksichtigung von Endogenität nicht zwangsläufig zu einer verbesserten Modellierung führt. Insgesamt kann die Studie als geeigneter Startpunkt für zukünftige Studien zu Bilanzpolitik im Bankensektor fungieren. Im anschließenden Kapitel 2.2 werden aus Gründen der Konsistenz Modellierung und Ergebnisse der Schätzung von Bilanzpolitik in nicht-finanziellen Industrien für die Studien in Kapitel 3 und 4 vorgestellt.

Im Rahmen einer umfassenden Betrachtung der Qualität von Finanzinformationen sollte die bereits angesprochene Verlässlichkeit der Rechnungslegung sichergestellt werden können. Hierbei spielt neben den Rahmenbedingungen der Rechnungslegung die Qualität der Attribute und Prozesse der Abschlussprüfung eine gewichtige Rolle. Im Zuge dessen wird die Unabhängigkeit des externen Abschlussprüfers von seinem Mandanten als eines der Kernattribute proklamiert, da diese Neutralität verleiht und simultan die Qualität der vollzogenen Abschlussprüfungshandlungen erhöht. Ein unabhängiger Wirtschaftsprüfer ist demzufolge besser in der Lage, das Ausmaß an Bilanzpolitik zu verringern, was wiederum eine Erhöhung der Qualität der Finanzinformationen nach sich zieht. Kapitel 3 widmet sich daher der Analyse der Unabhängigkeit von kleinen und mittelgroßen Abschlussprüfern im deutschen Markt für Mandanten von öffentlichem Interesse zwischen 2007 und 2014. Das Setting der Studie zeichnet sich dadurch aus, dass kleine und mittelgroße Wirtschaftsprüfer traditionell im Verdacht stehen, aufgrund einer eingeschränkten Unabhängigkeit von kapitalmarktorientierten Mandanten diesen ein höheres Maß an Spielräumen zuzugestehen, c.p., es wird eine geringere Qualität der Prüfung und damit

größerer Spielraum für Bilanzpolitik vermutet. Das europäische und insbesondere deutsche Umfeld qualifiziert sich durch divergente gesetzlich vorgeschriebene Schwellenwerte der Abhängigkeit des Wirtschaftsprüfers für eine differenzierte Analyse der Nichtlinearität der Zusammenhänge. Des Weiteren bilden die in Deutschland prävalenten Transparenzberichte eine profunde Datenbasis für die Ermittlung von Abhängigkeitsmaßen, die im speziellen Marktsegment eine überlegene Messung der Abhängigkeit der betrachteten Prüfer ermöglichen. Die vorliegende Arbeit in diesem Kapitel wählt auf Basis der prävalenten gesetzlichen Regelungen ein nicht-lineares Forschungsdesign und legt nahe, dass kleine und mittelgroße Wirtschaftsprüfers im speziellen Marktsegment der kapitalmarktorientierten Mandanten mit steigender Abhängigkeit eine erhöhte Prüfungsqualität liefern, bis sich die Effekte an einer hohen Schwelle der Abhängigkeit umkehren. Dieser Wendepunkt spiegelt zugleich formulierte gesetzliche Schwellenwerte wider. Kleinen und mittelgroße Abschlussprüfer scheinen trotz steigender Abhängigkeit eine höhere Prüfungsanstrengung zu wählen, welche buchmäßige Bilanzpolitik einschränkt und damit die Qualität der Finanzinformationen erhöht. Im Fokus steht hierbei vermutlich das Bestreben des Festigens und Ausbauens einer Reputation im Marktsegment der kapitalmarktorientierten Mandanten. Ab Überschreiten eines kritischen Schwellenwertes scheint hingegen die gestiegene Abhängigkeit die Reputationseffekte zu dominieren. Die Ergebnisse der Studie legen zudem den Schluss nahe, dass Mandanten bei einer Einschränkung der buchmäßigen Bilanzpolitik geneigt sind mehr reale Bilanzpolitik einzusetzen, deren Einschränkung nicht originärer Gegenstand der Prüfung durch den Abschlussprüfer ist. Die geschilderten Ergebnisse erweisen sich als sehr robust gegenüber zahlreichen Veränderungen und Tests wie etwa einer de-facto Office Level und Partner Level Analyse. Zudem kann gezeigt werden, dass die dominierenden Reputationseffekte wohl in der eingeschränkten Wichtigkeit der kapitalmarktorientierten Mandanten für den Gesamtumsatz begründet liegen, welche von bisherigen Studien systematisch überschätzt wird.

Kapitel 4 richtet den Fokus auf originär nicht-finanzielle Nachhaltigkeitsstrategien und deren Einfluss auf bilanzpolitisches Verhalten. Bisherige Studien betrachten die heterogene Gruppe der Corporate Social Responsibility (CSR) Unternehmen dabei überwiegend in ihrer Gesamtheit. Die vorliegende Analyse in Kapitel 4 versucht demgegenüber

die untersuchten Unternehmen in prävalente Anreizgruppen zu trennen, um vermutete diverse Anreize zur Investition in CSR mit den Anreizen zu Bilanzpolitik in Verbindung zu bringen. Hierfür wird ein europäisches Setting mit kapitalmarktorientierten Unternehmen der ASSET4 Datenbank zwischen 2005 und 2014 gewählt. Durch eine gezielte Verknüpfung der Anreize zur Investition mit der Entscheidung zur einschlägigen Berichterstattung werden Unternehmen mit prävalenten Reputationsanreizen von Unternehmen mit Tendenzen zur intrinsischen, philanthropischen Motivation abgegrenzt. Die Ergebnisse lassen vermuten, dass Unternehmen mit Reputationsanreizen bilanzpolitisch konform handeln und buchmäßige durch reale Bilanzpolitik zu ersetzen versuchen. Hierdurch lassen sich sowohl Risiken der medialen Bilanzschelte minimieren als auch hohe Gewinne als Teil einer Gesamtreputationsstrategie realisieren. Des Weiteren kann gezeigt werden, dass diese strategische Positionierung einen Einfluss auf die Wirkungsweise von Corporate Governance (CG) hat. Während Maße für gute CG mit einem geringeren Ausmaß an buchmäßiger Bilanzpolitik einhergehen, wählen Unternehmen mit Reputationsanreizen kohärent mit ihrer strategischen Tendenz erneut ein höheres Maß an realer Bilanzpolitik. Insgesamt deuten die Ergebnisse der Studie daraufhin, dass gewisse Unternehmen CSR als Reputationsinstrument nutzen und sich nicht konform einer nachhaltigen Strategie im Hinblick auf Bilanzpolitik verhalten.

Contents

List of tables

List of equations

List of figures

1 Introduction and summary

In an Arrow-Debreu world of unrestricted access to perfect and competitive financial markets, there is no need for accounting information about the financial situation of a firm. Because information is costless, share- and stakeholders are then indifferent in deposits and securities (e.g., Holthausen & Watts 2001; Freixas & Rochet 2008). However, several reasons exist indicating a rejection of the assumptions for an Arrow-Debreu world, hence there is no perfect costless information. Moreover, the distribution of information is asymmetric, causing follow-through multi-level agency problems, which are the main reasoning for the variety of financial and non-financial accounting standards, regulatory and advisory entities and the auditing and rating agency profession. Likewise, these agency problems have been at the heart of the accounting literature and raised the question of whether and how accounting information can help resolve these problems.

In their review on the earnings quality (EQ) literature, Dechow et al. (2010) provide an important framework for a general thinking about reporting with a focus on all kinds of financial information of a representative firm. Defining X as the unobservable performance of such a firm[1] and $f(\,)$ as the respective accounting system to map performance into observable accounting numbers, the innocent reader regards the following expositional proposition, presented in the form of an idealized formula, as straightforward:

$$Reported\ Earnings \equiv f(X)$$

When it comes to the quality of financial reporting, the earnings figure as the key accounting information should precisely capture the performance of the firm.

However, when thinking about the operationalization of this framework, several issues become apparent. The aggregation to a single number requires a sharp refinement and focus. Financial performance X could be simply defined as the sum of cash flows generated by the entity in the financial year. However, in a multi-year view of the performance of a firm, the financial performance might further subsume expected discounted cash-

[1] The specific element of performance information required is highly dependent on the type of stakeholder looking for performance information, e.g., the specific decision model (e.g., Dechow et al. 2010).

flows collocated within the future years of existence plus a possible change in the liquidation value of the firm (Sterling 1970; Ohlson 2000). While reducing the financial performance to one number X might be a conception of the reality that is too simple, **Figure 1** shows results from the survey by Graham et al. (2005) and highlights that CEOs/CFOs see accounting earnings as the most important performance number reported to outsiders (51 percent of the respondents). Hence, the informational content of a single earnings number should have superior informational content compared to simple cash flows.

Figure 1 – Most important performance measures from survey by Graham et al. (2005)

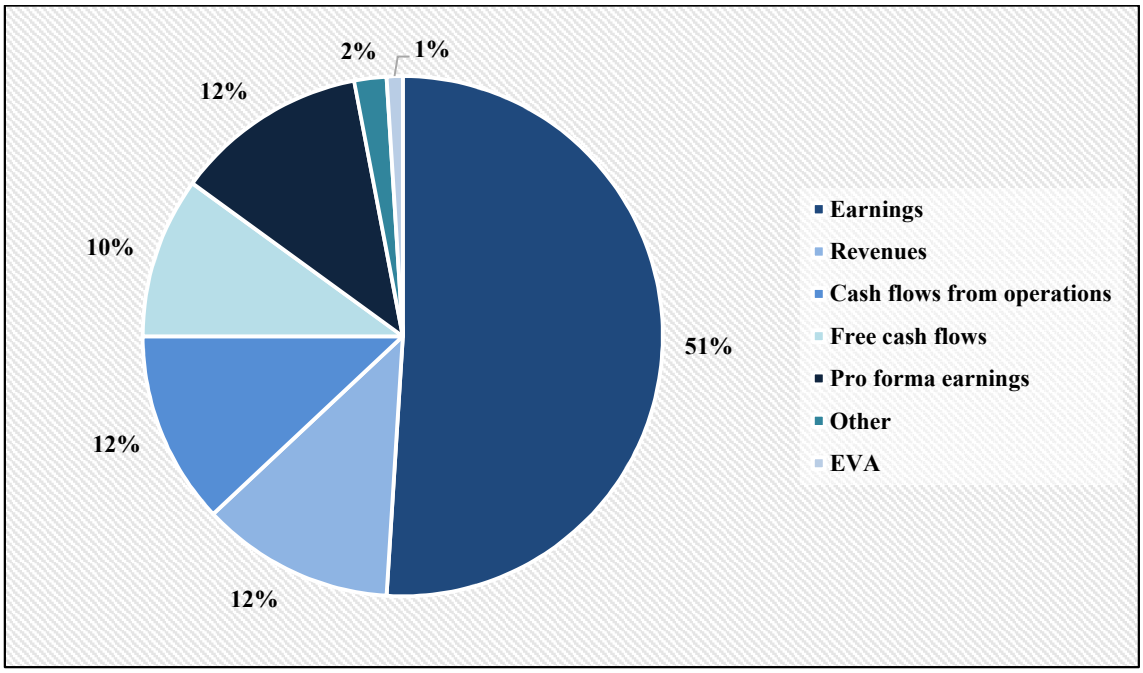


Figure shows results for 305 survey responses to the question: "Rank the three most important performance measures reported to outsiders." based on a survey of 401 financial executives. Figure shows percentages of respondents assigning #1 ranking to the respective performance. Source of information: Graham et al. (2005), p. 20.

Consequently, when thinking about the presentation of financial information, $f(\,)$ plays an important role.[2] Being a vehicle of communication, the management of the firm has incentives to see financial reporting not as the single compliance task towards investors, but the opportunity to address and influence investors, employees, suppliers, customers,

[2] In their extensive survey, Dichev et al. (2013) also find that CFOs regard accounting standards, in addition to the business model of the company, as the most important influencing factor of earnings quality.

lenders and the government. For investors, not only reliable judgement about the appropriateness of the work of the management, but also the depiction of performance and expectations about future opportunities and risks is inevitable.

International standard setting boards (e.g., FASB for the US-GAAP, IASB for the IFRS) have to balance rules and discretion in financial reporting standards to enable presentation of a true and fair view of a firm, thereby maximizing information efficiency of the financial report for multiple stakeholders with a variety of decision models (e.g., Kothari et al. 2010). Therefore, standard setters need to make trade-offs to enable standards to represent all users' needs in an appropriate way, which clearly leaves all addressees with suboptimal representation. On top of that, in the recent 20 years, complexity and multitude of business activities of globally active companies has likewise increased the complexity of the task to map all business activities from a variety of firms from different industries into one financial report that can feature value-enhancing attributes for addressees of the financial information. While this complexity has gradually led to an increased volume of financial reports (many financial reports account for more than 400 pages, e.g., KPMG 2014) and hence a possible information overload, the speed of global markets has raised the collective demand for more and more timely information, ideally with predictive value (CFA Institute 2013).

In his speech, IASB chair Hoogervorst (2017) talks about the complexity of the task to provide sufficient accounting systems:

> "Valuable information gets drowned out by 'tick the box' disclosures and voluminous, but poorly organised and presented, financial data. […] The fact is that IFRS Standards prescribe very little in the way of formatting the income statement. We define revenue, we define profit or loss, but we do not define very much in between." Hoogervorst (2017)

Ultimately, the variety of firms and information as well as the necessity of timeliness and prediction demand for high flexibility in assembling $f(\,)$, causing potential for intentional and unintentional errors and biases of information (e.g., Dechow et al. 2010). To frame the quality of accounting information in more detail, Gaynor et al. (2016) identify completeness (e.g., Botosan 1997; Botosan & Plumlee 2002), confirmatory value (Dechow &

Dichev 2002) and predictive value (e.g., Clor-Proell et al. 2014; Müller et al. 2015) as well as neutrality and freedom of error as main dimensions of quality. This doctoral thesis focuses on the last part of the picture, in particular on the intentional errors caused by management's discretion.

> "I fear that we are witnessing an erosion in the quality of earnings, and therefore, the quality of financial reporting. Managing may be giving way to manipulation. Integrity may be losing out to illusion." (Levitt 1998)

Levitt's famous speech effectively raised awareness for the flexibility in accounting being a way for managers to manipulate earnings numbers for their own good. Consequently, it has caused major changes in both the accounting and auditing profession to ensure the validity of accounting information (e.g., Beatty & Liao 2014; DeFond & Zhang 2014). Nevertheless, intentional errors, hereinafter referred to as Earnings Management (EM), is a phenomenon that is likely to prevail as long as accounting information seeks to map a true and fair view of the firm to eliminate information asymmetries using expectations, estimates and subjective opinions by the management, who possess information advantages. Dichev et al. (2013) find in their recent survey that EM is modestly used by top management, e.g., based on their survey 20 % of the firms use 10 % of their earnings representing 2 % economy-wide EM. Moreover, accounting research has also highlighted the possibility of managers to use real actions to manage earnings in a certain way. In their survey, Graham et al. (2005) highlight (see **Figure 2**) that 80% of the survey participants would decrease discretionary expenses to manage earnings upwards, while this is an alteration of real business actions resulting in higher earnings. However, the accounting of the discretionary expense information is not influenced by this type of discretion. Subsequently, there have been many studies (e.g., Roychowdhury 2006; Cohen et al. 2008; Cohen & Zarowin 2010; Gunny 2010) focusing on real activities earnings management (REM) in addition to the originally studied accounting based accrual earnings management (AEM), in particular with respect to a possible trade-off relation between the two types of EM (e.g., Zang 2012).

Figure 2 – Importance of REM vs. AEM from survey by Graham et al. (2005)

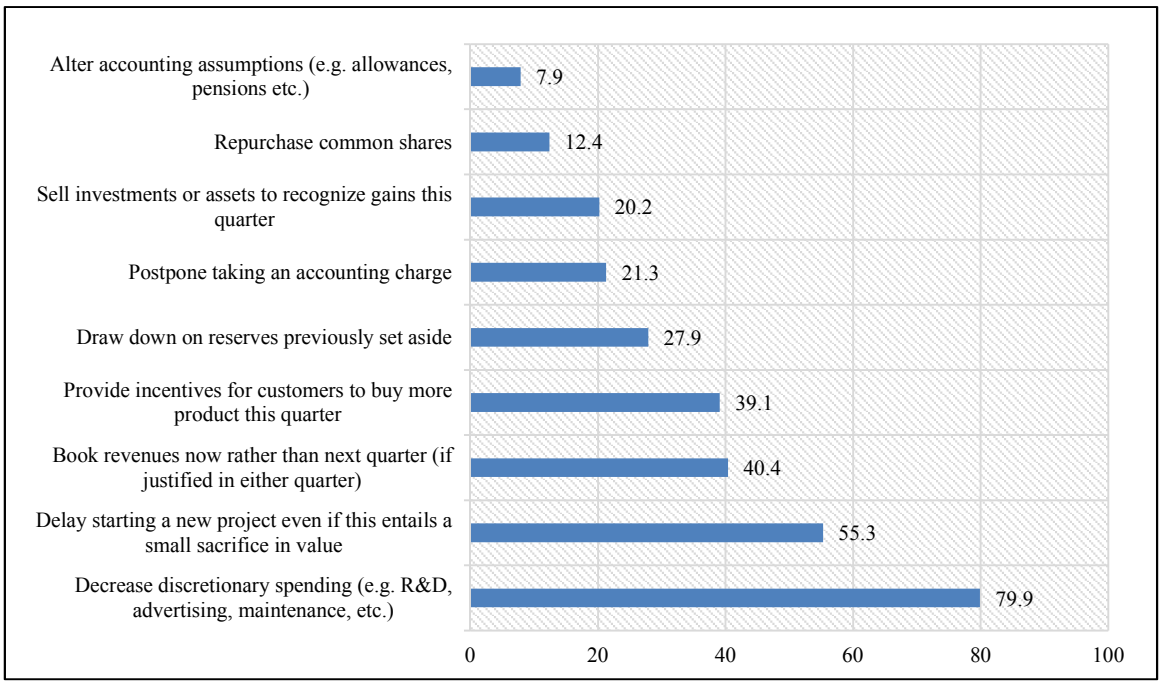


Figure shows results for 298-304 survey responses to the question: "Near the end of the quarter, it looks like your company might come in below the desired earnings target. Within what is permitted by GAAP, which of the following choices might your company make?" based on a survey of 401 financial executives. Numbers show percentages of respondents agreeing or strongly agreeing with the respective statement. Source of information: Graham et al. (2005), p. 33-35.

To be able to study the phenomenon of EM in empirical settings throughout this doctoral thesis, chapter 2 deals with the comprehension of appropriate proxies to measure both kinds of EM. Accounting research has traditionally studied the field of EM separately for non-financial and financial industries, since financial institutions have disparate balance sheets and diverse incentives to pursue EM compared to non-financial firms. Since EM cannot be observed directly, it is important for every research question in any setting to find a verifiable proxy for EM. What is important, in the light of the earlier discussion about asymmetric information and follow-up agency problems, banks as financial intermediaries function as a vehicle to address information asymmetry between borrowers and depositors. They enhance much more efficient (because of high economies of scale) and less costly (because of no redundant tasks) monitoring of the borrower compared to the vast number households (Diamond 1984).[3] Decisively, this upside comes at the cost of a potential information asymmetry between depositors and bank managers, which can

[3] See Armstrong et al. (2010) for a thorough discussion of how bank financing can help mitigate information asymmetries between borrowers and lenders.

cause incentives for intentional management of accounting numbers on another level. While numerous studies on AEM in banks exist, accounting research still lacks a thorough understanding of how to measure AEM, e.g., what regressors can add value to the estimation process. The analysis in chapter 2.1 tries to close this gap and analyses existing model specifications of discretionary loan loss provisions (LLP) in the banking sector to identify common pattern groups and specific patterns used. Thereupon, prevalent test procedures are applied to an U.S.-dataset from 2005-2015 to examine the extent of measurement errors, extreme performance and omitted-variable biases as well as predictive power of the discretionary proxies of each of the models. The results highlight the importance of a thorough understanding about the methodological modelling process of EM in the banking industry. The currently established models to estimate EM are appropriate yet optimizable. In particular, the analysis identifies non-performing asset patterns as the most important group, while loan loss allowances and net charge offs can add some value, though do not seem to be indispensable. In addition, the results show that non-linearity of certain regressors could be an issue, which should be addressed in future research, while the analysis identifies some omitted and possibly correlated variables that might add value to specifications in identifying non-discretionary LLP. Results also indicate that a dynamic model and endogeneity robust estimation approach is not necessarily linked to better prediction power. To enhance the coherence of the analysis, chapter 2.2 depicts commonly applied models of the industry part of the accounting literature on AEM and REM and presents the results from first stage regressions of the models applied in chapter 3 and chapter 4.

Chapter 3 sheds the light on another important vehicle of improvement of financial reporting quality, e.g., the influence of external audits with different quality. Therefore, the study undertakes an in-depth analysis of one important input factor of the auditing process and audit quality, which shapes the quality of the financial information provided by the company. In particular, DeFond and Zhang (2014) and Gaynor et al. (2016) highlight the advantage of high quality audits when the quality of financial information is low, e.g., a high quality audit can restrict the amount of EM used by the management. The analysis in chapter 3 uses a German sample of clients from small and midsized auditors with 1,052

firm-year observations between 2007 and 2014 and studies the influence of client importance, e.g., economic dependence on the client, on the amount of EM used. In particular, the study consciously concentrates on a market (German small and midsized auditors) where client importance is claimed to be more pronounced than in markets with high concentrations of large Big N auditors. What is more, the setting is special with regard to existent regulatory thresholds of economic dependence as well as opposing incentives and very mixed results, indicating a non-linearity in the relationship that has not been studied thoroughly. Consequently, the study raises the question whether firms audited by small and midsized audit companies use higher AEM around certain critical client importance thresholds because their audit quality is impaired. Furthermore the analysis looks at whether these firms trade-off REM as their use of AEM increases/decreases. Two precise client importance proxies are investigated, (1) measured as total sales from the client in relation to total sales of the audit firm from all clients and (2) measured as non-audit sales in relation to total sales from the client. By using (1), client importance can be measured much more precisely than in the most part of prior literature. The results indicate for measure (1) that firms with auditor's client importance below the thresholds generally use lower amounts of AEM and trade-off from AEM to REM as client importance increases. Hence, there is an increase in audit quality, presumably caused by superimposing reputation and litigation incentives of the auditors in this special market segment. However, for firms with auditor's client importance above 5%, 10% and 15% for one, two and three consecutive years, we find that AEM increases and REM decreases again as client importance increases, while the intensity of the effects increases with an increasing time-period and magnitude of the threshold. This can be seen as a sign for the dominance of diverse incentives at different levels of dependence. In contrast, we do not find significant results for measure (2), indicating the necessity to study non-audit service importance and total sales client importance separately and precisely. The results hold for de-facto office and partner level analysis, while a convincing reasoning is provided, which opens up a new perspective on small and midsized auditors and their eligibility to perform high quality PIE audits. What is more, the analysis tries to cancel out the moderating role of audit committees in changing or causing the results, while providing some interesting insights into the effectiveness of supervisory boards and the additional benefit

of a best practice audit committee. Furthermore, the various other alterations show the robustness of the findings, e.g., using alternative EM proxies, using additional confounding controls as well as various changes in methodology.

Chapter 4 connects the fields of financial information and non-financial information on Corporate Social Responsibility (CSR) activities in a setting where there is no mandatory framework for non-financial reporting, but high incentives to act socially responsible for different reasons. The study tries to elaborate diverse incentives to perform well in CSR, while one of these might be connected to opportunistic EM behaviour as part of a reputation enhancement strategy. In particular, the analysis in chapter 4 examines whether firms that act socially responsible engage in different amounts of EM. Therefore, this piece of work tries to disentangle the two main incentives incorporated into CSR investments and analyses whether these entail diverse EM strategies in terms of AEM, REM, total EM and the trade-off relationship between AEM and REM. Furthermore, the effect of Corporate Governance (CG) on EM strategies is considered, particularly whether the respective CSR incentives moderate this relationship. For this purpose, a European sample with 2,733 firm-years from the 2005-2014 period is comprehended. The results show that EM strategies might vary between intrinsic and reputational CSR incentives. For both incentives, the respective firms seem to engage in lower AEM. However, high-reputation firms use more total EM and trade off from AEM to REM at the possible expense of shareholder value, whereas rather intrinsic firms use less total EM and trade off vice versa. Additionally, the results reveal that certain independence in the board limits opportunities to pursue AEM, whereas CSR incentives significantly moderate the outcome of effective monitoring. While both CSR incentives entail a lower use of AEM with increasing independence, high-reputation firms again switch to higher REM and further trade off from AEM to REM. Altogether, the results highlight that firms with a high CSR orientation engage in different EM strategies, depending on their CSR incentives and thereby presumably preserve or jeopardize shareholder value.

2 Earnings management modelling in accounting research

To be able to study the field of EM properly, research studies have to identify appropriate EM proxies. While cash flows are regarded as a noisy proxy for performance, accruals function as an instrument to produce a decision-useful proxy for financial performance, hence accounting earnings (e.g., Dechow 1994; Dechow & Dichev 2002). As discussed in chapter 1, the accounting system provides a way of reporting an earnings number capturing a true and fair view of the financial situation of the firm, while discretion is provided to each of the firms to meet this requirement. When measuring EM, most studies rely on the total amount of accruals as an initial measurement of the total adjustment of the cash-flow numbers. When it comes to modelling the accrual part that is likely to be affected by management discretion, there are several studies that discuss the statistical issues in empirical accrual modelling (e.g., Dechow et al. 1995; McNichols & Wilson 1988; McNichols 2000). They partition the total amount of accruals into a non-discretionary (NA) part caused by normal developments within the firm, and a discretionary (DA) component, which is assumed to be related to EM behaviour:

Equation 1

$$TA_{it} = DA\_true_{it} + NA\_true_{it}$$

Ideally, the applied models would capture the true amount of normal variation (NA_true_{it}) of the total accruals to comprehend the remaining part as the true discretionary variation (DA_true_{it}). Since both parts cannot be observed, research uses a simple linear panel regression framework with N individuals and K regressors to comprehend a proxy DA_{it}:

Equation 2

$$TA_{it} = \alpha_0 + \boldsymbol{\beta}\, \boldsymbol{NAX}_{it} + u_{it}$$

TA_{it} is the total accruals of firm i in year t, $\boldsymbol{NAX}_{it}$ is a $N \times K$ matrix of certain variables that are assumed to cause variation, which is connected to non-discretionary business

decisions and developments. DA_{it} is calculated as the residual u_{it} of **Equation 2**, with a certain measurement error η_{it}:

Equation 3

$$DA_{it} = DA\_true_{it} + \eta_{it}$$

The properties of η_{it} determine whether DA_{it} is a good or a noisy measure. These properties are dependent on the accuracy of our non-discretionary regressors in estimating the non-discretionary variation, e.g., NA_true_{it}:

Equation 4

$$\eta_{it} = NA_{it} - NA\_true_{it}$$

Consequently, research studies use the following simple linear regression:

Equation 5

$$DA_{it} = \alpha_0 + \beta_1 VOI_{it} + \boldsymbol{\beta}\, \boldsymbol{X}_{it} + \varepsilon_{it}$$

VOI_{it} is a dichotomous or continuous variable of interest that is hypothesized to exert significant influence on discretionary accruals, e.g., EM. $\boldsymbol{X}_{it}$ is another $N \times K$ matrix that captures a set of control variables that influences the variation in the discretionary dependent variable, while ε_{it} is an error term capturing the individual heterogeneity of the firm i in year t.

If the coefficient for the variable of interest (β_1) has the hypothesised sign and is statistically significant at the conventional (0.01, 0.05, 0.1) levels, researchers consider their results to support their hypotheses. In this situation, the significance of coefficient β_1 can be biased when the variable of interest is correlated with the error η_{it}. In particular, this will be the case when NA_true_{it} is correlated with VOI_{it} but NA_{it} does not capture NA_true_{it} properly, e.g., η_{it} is not just white noise.

The remainder of this chapter focuses on the special setting of banks in chapter 2.1 and presents an extensive analysis of the applied models to reveal information about the validity of bank accounting studies when it comes to measuring EM. In particular, the studies tries to shed light on the accuracy of fit of the specifications used in disentangling NA_true_{it} and therefore limiting η_{it} to white noise. In chapter 2.2, results of modelling

analyses in the non-financial part of the accounting literature are briefly discussed, whereinafter modelling approach and results of first-stage regressions for the empirical analyses in chapters 3 & 4 are presented.

2.1 Earnings management modelling in the banking industry[4]

2.1.1 Introduction

Earnings management (EM) is a frequently discussed topic in both accounting and banking literature. This fact is not surprising, since EM is a very debatable, sometimes even obstructive issue when it comes to discussing earnings quality and the need for relevant and reliable information disclosure in its entirety (e.g., Ball & Shivakumar 2005).

> "Higher quality earnings provide more information about the features of a firm's financial performance that are relevant to a specific decision made by a specific decision-maker." (Dechow et al. 2010)

In this mind-set, research settings typically capture an opportunistic use of EM to manipulate the perception of the state of a company in favour of the management and/or company itself. Over the last three decades, various studies in the accounting and banking field have proposed models, which are able to estimate the non-discretionary variation of total or parts of total accruals. Based on these models, they isolate the discretionary part of the total accruals, which can be seen as a potential proxy for accounting-based EM. However, since a direct measurement is not possible,[5] these modelling attempts have considerably changed over time, with the goal to improve preciseness and credibility.[6]

Altogether, EM models in the accounting literature are characterized by two main properties. First, they approach the measurement from various directions, all of which are justifiable and appropriately reasoned. Second, a clear path of improvement is discernible, as the motivation for new models is connected with the flaws of the previous ones. Both

[4] This chapter is based on a 2017 version of a working paper titled "Earnings management modelling in the banking industry – evaluating valuable approaches", which is co-authored by Markus Stralla. The reasoning, results, and interpretations of this might change after the submission of this thesis and the completion of the doctoral degree. The most recent version of this study is available upon request. Please do not cite this working paper without permission.
[5] Although a direct measurement is not possible, there are surveys investigating the existence and motivation behind EM, e.g. Graham et al. (2005). In spite of that, general possibility to pursue EM likely is likely to cause a certain utilization.
[6] See chapter 2.2 and/or Dechow et al. (2010) for a general overview on the modelling of discretionary accruals in the other industries.

characteristics can be transferred to the EM modelling approach in the banking literature. Beatty and Liao (2014) compare various EM models and study these two properties in their comprehensive literature review. They discuss the first main difference of EM measurement in the banking industry, which is the focus on a single accrual, e.g., loan loss provisions (LLP). The authors provide three potential explanations for this focus. First, LLP is by far the most important accrual, amounting to 56% of total accruals and 34% of total accrual variance. Second, the focus on LLP could be related to the urge of minimizing measurement error. When total accruals are considered, mapping non-discretionary variation might be more prone to measurement issues. Third, data availability may also play a role, since major databanks gradually developed and provided all variables for more complex approaches. The aforementioned reasons are plausible, which is why we focus on LLP models in our extensive analysis of previous studies published in highly ranked accounting and banking journals. The second main difference between EM in accounting and banking is the missing comprehensive discussion and examination of the applied models. Beatty and Liao (2014) state that there is no "consensus in banking studies on how to best model discretionary accruals." What is important, existing models vary considerably in complexity and choice of pattern groups. In addition, in contrast to the accounting literature for the non-financial industries (e.g., Dechow et al. 1995; Young 1999; Peasnell et al. 2000), there are no studies on the modelling approach for banks, which is why we try to provide important insights to fill this gap.

Our results indicate that non-performing loan patterns are the most important pattern group when it comes to separating LLP into discretionary and non-discretionary parts, while loan loss reserve and/or net charge-off patterns can enhance the modelling, though are less important. Consequently, EM proxies derived from models that are more complex predict potential EM more accurately, while for settings of data limitation, sufficient models can be found. Furthermore, we find that the relationship between LLP and a common set of control variables might be non-linear in contrast to common assumptions of linearity, while we identify growth, loan intensity, income diversification and operating cashflow patterns as possible omitted correlated variables that could improve the quality of discretionary LLP. At last, we find that dynamic modelling seems to improve explanatory power, while they might explain some discretionary part of total LLP, which can lead to

biased inferences and lower prediction power. We also show that endogeneity robust estimation cannot solve this problem.

Altogether, we contribute to the current literature in two major ways. First, we provide an overview of the differences between various EM estimation procedures and analyse the respective regression pattern groups. Furthermore, we extent our research on models for situations of data availability issues, which are especially relevant for research settings that investigate non-commercial banks. Second, we apply several test procedures from banking and accounting literature to investigate measurement errors, omitted variable biases and prediction power of the applied EM models. Based on our results, future research should be able to identify an appropriate specification and further improve the modelling of EM to enhance the validity of inferences drawn from regressions.

The remainder of the paper is organised as follows. Section 2.1.2 presents a brief overview of EM modelling in the banking literature and provides an extensive analysis of the specifications used in prior literature, whereupon we set up various models for our test procedures. Section 2.1.3 contains the research design, including data, sample selection and empirical test procedures to test the validity of the models. Section 2.1.4 presents the results from first-stage regressions and the uni- and multivariate findings of the empirical test procedures. Section 2.1.5 presents a summary and conclusion.

2.1.2 Earnings management in banks and modelling discretionary LLP

2.1.2.1 Earnings management in banks

Major parts of the earnings and financial reporting quality literature focus on evaluating the degree of EM in non-financial firm samples.[7] Consecutively, EM in banks has evolved as a somewhat distinct field of research, particularly because possibilities and incentives to pursue EM differ significantly and therefore determine customized methods of measurement.

Beatty and Liao (2014) provide an extensive theoretical and empirical overview of banking research in accounting. Among various related questions, they elaborate and discuss

[7] Dechow et al. (2010) give an extensive literature overview for the quality of earnings research. When discussing EM, they also focus on the measures that are used in studies that cover non-financial industry questions.

why banks use discretion in their financial reporting to foster certain opportunistic goals, namely capital management (e.g., Wahlen 1994; Collins et al. 1995; Beatty et al. 1995; Kim & Kross 1998) and smoothing of earnings (e.g., Ahmed et al. 1999; Anandarajan et al. 2007; Beatty & Liao 2009; Bouvatier et al. 2014). Bushman (2014) complements Beatty and Liao (2014) and highlights the risk-taking aspect of opportunistic reporting behaviour. Acharya and Ryan (2016) extensively discuss selected papers and provide suggestions for regulation to enhance the stability of the financial system.

These studies highlight the decisiveness of discretionary accounting choices in banks and hence the importance of research in this field. Essentially, for adequate inferences to be drawn, underlying methodological approaches have to be adequate. Beatty and Liao (2014) give an idea of how many different measures exist, while actively stating the importance of a thorough analysis of the applied models and the modelling approach itself. Although applying significantly different models, most studies use some sort of discretionary accruals, which is equivalent to the non-financial part of the accounting literature. Models vary over time in the analysed dependent variables. For example, some studies use discretionary loan loss reserves (e.g., Hasan & Wall 2004; Jin et al. 2016) or realized gains and losses (e.g., Beatty et al. 1995; Collins et al. 1995; Beatty & Harris 1998; Beatty et al. 2002). Yet, most papers regard LLP as the most important vehicle for EM in banks (e.g., Bushman 2014; Lobo 2017). LLP account for 56% of the total accruals (e.g., Beatty & Liao 2014) as well as 15-20% of earnings before taxes and loan loss provisioning ($EBTP$) (e.g., Lobo 2017). In particular, a bank using higher LLP can intentionally build up loan loss reserves in years of high performance for means of improving earnings numbers when $EBTP$ is low (e.g., Sutton 1997; Levitt 1998). Alternatively, models could follow a total accrual approach, equivalent to the industry models. However, several studies (e.g., Dechow et al. 1995; Peasnell et al. 2000; McNichols 2002; Dechow et al. 2010) already show that total accrual designs are associated with potentially high degrees of measurement error in non-financial settings, particularly due to omitted correlated variables and the complexity of identifying appropriate normal accrual regressors. What is even more important, for financial entities, the remaining 44% of the total accruals are unlikely to be subject to discretion since standard setting for banks has consequently limited accounting flexibilities (e.g., Beatty & Liao 2014). Hence, total accrual approaches

would be likely to produce additional measurement error rather than feasible discretionary accruals.

Furthermore, the degree of discretion in the provisioning for loan losses has likewise been shaped by significant changes in accounting standards for credit losses. While the respective rules in the 1990s relied heavily on future-oriented fundamentals, e.g., non-performing loans, to evaluate their loan loss reserves (e.g., Ludwig 2009; Beck & Narayanamoorthy 2013; Beatty & Liao 2014), the SEC and FASB issued the Staff Accounting Bulletin (SAB) 102 and the Federal Financial Institutions Examination Council (FFIEC) Policy Statement (e.g., FFIEC 2001; SEC 2001) for fiscal years after 2001. These emphasized a market-to-market based evaluation in the form of the incurred loss model, e.g., focus on already occurred loss events for means of loan loss provisioning, resulting in less room for discretion. Subsequently, there has been a constant field of tension between the decision usefulness of accounting standards and the regulatory prevention of bank failures, particularly due to the pro-cyclicality of provisioning for loans and leases under the incurred loss model (e.g., Financial Stability Forum 2009; Bushman & Williams 2012). As a result, there have been changes in LLP accounting standards once again, e.g., the introduction of the expected loss model in IFRS 9 and the current expected credit loss model in ASC 326. They leave banks again with more discretion in terms of the inclusion of future estimates (e.g., Bushman & Williams 2015; Lobo 2017; PwC 2017) for fiscal years 2018 (IFRS 9) and 2020 (ASC 326).[8] Altogether, LLP are likely to remain the most important discretionary accrual for banks, which is why we concentrate on the discretionary LLP models throughout our analysis.

2.1.2.2 Specification analysis

We focus on a sample of studies from the accounting and finance literature between 1990 and 2017. We incorporate all regression-based analyses published in journals with an h5-index above 50 or an SRJ above 0.500.[9] We identify 39 papers using discretionary LLP

[8] While, neither framework requires a specific methodology, IFRS 9 and ASC 326 require "the estimate of expected credit losses […] [to] consider historical information (past events), information about current conditions, and reasonable and supportable forecasts of future events and economic conditions, as well as estimates of prepayments." (PwC 2017, p. 6) In addition, US-GAAP does not require multiple forward-looking scenarios as long as the scenario is carefully selected and represents the expected credit loss. Altogether, US-GAAP and IFRS guidelines seem to provide room for estimates about the future and therefore discretion.

[9] The analysed papers use one- and two-step approaches, where one-step means inclusion of variables of interest in the first-stage regressions together with the non-discretionary LLP regressors, while two-step means separate, second-stage analyses of variables

as their EM or EQ measure and capture time-period analysed, setting, one vs. two-step modelling and specifications.[10] We identify eight pattern groups in the literature.[11] A brief look at panel A of **Table 1** shows the frequencies of use for all pattern groups as well as the specific patterns from each group. Most papers (76.9 percent) use both total loan and non-performing loan patterns to estimate non-discretionary LLP, whereas 48.7 percent respectively 43.6 percent of the studies use loan loss reserves respectively net loan charge-offs. In addition, a further set of controls is included by 59.0 percent of the studies for capital requirement ratios and earnings before provisioning, while 35.9 percent use a variable for the overall size of the bank.[12] We separately study the total loan and non-performing loan pattern groups in panel B & C of **Table 1**, since most studies use more than one regressor to capture these pattern groups. Again, we find a variety of patterns for both groups, highlighting the importance of a structured and thorough analysis. For non-performing loans, most studies use a combination of NPL_{it-1} and ΔNPL_{it} (33.3 percent), while for the total loans group, ΔTL_{it} is the predominantly used (36.7 percent) pattern.

In the following, we comment on every pattern group and present respective reasoning for the choice of regressors. We further discuss the model specification parts and gradually develop a specification for our further analysis.

Total loans

Kanagaretnam et al. (2010a) elaborate that a higher level of loans results in a higher level of provisions (e.g., also Kim & Kross 1998). Therefore, they expect a positive relation. However, Bikker and Metzemakers (2005) argue that this relationship would only be true in a world of prudent and forward-looking banks, which is not in line with a reality where banks exaggerate expectations to minimize provisioning. To our

of interest with discretionary LLP as a dependent variable. We focus on a two-step approach to enable our empirical test procedures and further comment on this design choice whenever applicable.

[10] See Appendix B for the list of specifications.

[11] We do not report frequencies for lagged LLP as a possible pattern group, since we regard the decision to include these variables as a methodological one (static vs. dynamic modelling) and not a choice of patterns. Bouvatier et al. (2014) use LLP_{it-1} and Fonseca and Gonzalez (2008) additionally use LLP_{it-2} to implement a dynamic approach. We further discuss static vs. dynamic modelling in section 2.1.2.4.

[12] We find 23 papers with a one-step approach, while 16 papers comprehend discretionary LLP as the residual from a first stage model and test their variables of interest in a separate second stage. In addition, there is no change over time in the preference of one- or two-step approaches. We note that not all papers seek to draw inferences about EM behaviour, e.g., Liu and Ryan (1995) study the influence of bank loan-portfolio composition on loan loss provisioning. Rather, all papers seek to explain the variation of non-discretionary LLP.

understanding, the expected sign is unpredictable due to uncertainty about the incremental quality of the loan portfolio. What is more, changes in total loans can be seen as changing assessment of future default risk (e.g., Lobo & Yang 2001; Bikker & Metzemakers 2005), hence an increase would go hand in hand with an increase in provisioning (expected sign positive).[13] While 36.7 percent of all papers with a total loans pattern choose to include only ΔTL_{it}, only 13.3 percent apply TL_{it} and ΔTL_{it} (see panel C of **Table 1**). However, we choose to follow the latter approach, since we expect changes in assessment of future credit risk and level of actual credit risk to have distinct influences on non-discretionary LLP.[14]

Non-performing assets/loans[15]

Loans, whose payments are due for more than 90 days, are categorised as non-performing. Most studies include non-performing assets, since a higher level of non-performing assets signals difficulties with the loan portfolio (e.g., Wahlen 1994; Kanagaretnam et al. 2010a; Hamadi et al. 2016). These might force banks to act, resulting in higher provisions. This argumentation holds true for the change in non-performing assets. Increases reflect actual changes in default rate, e.g., improving or deteriorating loan portfolio quality (e.g., Beatty et al. 1995; Kim & Kross 1998; DeBoskey & Jiang 2012). Prior studies use patterns for level and change in non-performing assets with different lags. Based on our further analysis of the pattern group in panel B of **Table 1**, we add the pattern with the highest frequency of use, NPA_{it-1} and ΔNPA_{it} (33.3 percent of all NPA papers and 25.6 of all papers), to our specification pattern list.[16] The expected signs for the two coefficients are positive.

[13] Some studies note that the effect of changes in loans could also be ambiguous if changes in total loans are not caused by changes in credit default risk. Then, the effect of changes in loans might be ambiguous, just like the effect of total loans, because of uncertainty about the quality of the loan portfolio (e.g., Kanagaretnam et al. 2004; Hamadi et al. 2016).
[14] Kim and Kross (1998) also formulate $LLP_{it} = f(TL_{it}, TL_{it-1}, NPA_{it}, NPA_{it-1}, NCO_{it})$, while including change and level of total loans implicitly maps lagged total loans.
[15] Throughout our study, the two terms non-performing loans and non-performing assets are used interchangeably. The same holds for the two terms loan loss allowances and loan loss reserves.
[16] Since $NPA_{it-1} + \Delta NPA_{it} = NPA_{it}$, this pattern simultaneously captures the influence of current non-performing loans, which is the pattern with second highest frequency of use.

Set of further controls (capital adequacy ratio, earnings before provisioning and bank size)

Papers using one- and two-step approaches apply a multitude of confounding control variables to cancel out omitted variable biases. One component is the inclusion of important capital ratios to control for banks' incentives to manage their capital adequacy (e.g., Beatty et al. 1995; Ahmed et al. 1999; Bouvatier et al. 2014). Furthermore, prior research controls or tests for the earnings smoothing incentives of banks.

Table 1 – Specification pattern distribution

Specification pattern distribution					
Panel A: All regressors					
Pattern group	**Pattern**	**Number of pattern papers**	**(percentage of pattern group papers)**	**Number of pattern group papers**	**(percentage of papers)**
Total loans	ΔTL_{it}	21	(70.0%)	30	(76.9%)
	TL_{it}	12	(40.0%)		
	$TL\_CATEGORIES_{it}$	9	(30.0%)		
	ΔTL_{it+1}	1	(3.3%)		
	TL_{it-1}	1	(3.3%)		
Non-performing assets/loans	ΔNPA_{it}	22	(73.3%)	30	(76.9%)
	NPA_{it-1}	13	(43.3%)		
	NPA_{it}	9	(30.0%)		
	ΔNPA_{it+1}	6	(20.0%)		
	ΔNPA_{it-1}	3	(10.0%)		
	ΔNPA_{it-2}	2	(6.7%)		
	NPA_{it-2}	2	(6.7%)		
	ΔNPA_{it-3}	1	(3.3%)		
Capital requirement ratio	$CAPB_{it}$	16	(94.1%)	23	(59.0%)
	$CAPB_{it-1}$	7	(5.9%)		
Earnings before provisioning	$EBTP_{it}$	23	(100.0%)	23	(59.0%)
	$\Delta EBTP_{it}$	1	(4.3%)		
	$\Delta EBTP_{it+1}$	1	(4.3%)		
Loan loss allowances/reserves	ALW_{it-1}	13	(68.4%)	19	(48.7%)
	ALW_{it}	6	(31.6%)		
Net loan charge-offs	NCO_{it}	16	(94.1%)	17	(43.6%)
	ΔNCO_{it}	1	(5.9%)		
Bank size	$SIZE_{it}$	12	(85.7%)	14	(35.9%)
	$SIZE_{it-1}$	2	(14.3%)		
	$\Delta SIZE_{it}$	1	(7.1%)		
Macroeconomic variables (ΔGDP_{jt}, $\Delta UNEMP_{jt}$, $\Delta LandPrice_{jt}$, $CSRET_{jt}$, ΔBFI_{jt}, ΔSDA_{jt})				13	(33.3%)

Table 1 (continued)

Specification pattern distribution

Panel B: Non-performing asset/loan regressors		
Pattern	**Number of *NPA* papers**	**(percentage of *NPA* papers)**
$NPA_{it-1}\ \Delta NPA_{it}$	10	(33.3%)
NPA_{it}	5	(16.7%)
ΔNPA_{it}	4	(13.3%)
$NPA_{it}\ \Delta NPA_{it}$	3	(10.0%)
$NPA_{it-2}\ NPA_{it-1}\ \Delta NPA_{it}\ \Delta NPA_{it+1}$	2	(6.7%)
$NPA_{it}\ \Delta NPA_{it}$	1	(3.3%)
$\Delta NPA_{it-2}\ \Delta NPA_{it-1}\ \Delta NPA_{it}\ \Delta NPA_{it+1}$	1	(3.3%)
$\Delta NPA_{it}\ \Delta NPA_{it+1}$	1	(3.3%)
$NPA_{it-1}\ \Delta NPA_{it-1}$	1	(3.3%)
$\Delta NPA_{it-3}\ \Delta NPA_{it-2}\ \Delta NPA_{it-1}\ \Delta NPA_{it}$	1	(3.3%)
ΔNPA_{it+1}	1	(3.3%)
Panel C: Total loan regressors		
Pattern	**Number of *TL* papers**	**(percentage of *TL* papers)**
ΔTL_{it}	11	(36.7%)
TL_{it}	4	(13.3%)
$TL_{it}\ \Delta TL_{it}\ TL\_CATEGORIES_{it}$	4	(13.3%)
$TL_{it}\ \Delta TL_{it}$	4	(13.3%)
$TL\_CATEGORIES_{it}$	4	(13.3%)
$\Delta TL_{it}\ TL\_CATEGORIES_{it}$	1	(3.3%)
$TL_{it-1}\ \Delta TL_{it}$	1	(3.3%)
ΔTL_{it+1}	1	(3.3%)

Pattern distributions are based on the loan loss provision specifications of the papers stated in Appendix B. All variables are defined in Appendix A.

For the industry models, Kothari et al. (2005) highlight that without appropriate control for performance, the resulting EM proxies are biased, e.g., performance of the firm significantly influences the magnitude of the non-discretionary accruals and therefore exclusion leads to a bias of the discretionary accruals. Since most banking models focus on LLP single accrual approaches, they apply performance before provisioning (e.g., Ahmed et al. 1999; Anandarajan et al. 2007; Leventis et al. 2011; Bushman & Williams 2015).[17] Panel A in **Table 1** highlights that 59 percent of the analysed papers control for capital requirement ratios and earnings before provisioning, while the overwhelming majority use current year proxies.[18] We follow this approach and add $EBTP_{it}$ and $CAPB_{it}$ to our specification pattern list. As a third major control, 35.9 percent of the papers control for bank specific size (e.g., Agarwal et al.

[17] Kothari et al. (2005) distinguish between the inclusion of a performance control and actual matching. We assume actual matching to over-correct for performance, especially for discretionary LLP, since they are recurrent and systematic. E.g., a bank has positive discretionary LLP due to high EM and the matched bank has comparably high positive discretionary LLP, not due to bad fitting of the model, but because both banks pursue EM, e.g., have comparable earnings smoothing or capital management incentives. Performance matching would correct for actual EM and hence distort the results on the null hypothesis of no EM in a way that it accepts the null. We therefore only use $EBTP$ as an additional regressor.

[18] We note that capital adequacy ratios are applied in a wide variety, e.g., tier 1 capital to risk-weighted assets, tier 1 capital to total assets, tier 1 and tier 2 capital ratios, etc. We test for the validity of our results when changing the capital adequacy proxy used. However, our results are not dependent on this design choice.

2007; Beatty & Liao 2011; Beck & Narajanamoorthy 2013). Again, we decide upon the pattern option with the highest frequency of use (85.7 percent of the papers with a size pattern), e.g., $SIZE_{it}$.

Loan loss allowances/reserves

Allowances for loan losses capture past decisions about loan loss provisioning. The effect on current LLP and therefore its sign depends on the relation between past provisioning, the actual demand for provisioning and loan loss recognition (e.g., Beatty et al. 1995; Kanagaretnam et al. 2010a; DeBoskey & Jiang 2012). If larger past provisioning, summarized by loan loss reserves, is associated with increases in loss recognition, especially in times of distrust and high uncertainty, larger reserves should be associated with larger current loan loss provisions, hence we expect a positive sign (e.g., Wahlen 1994; Ahmed et al. 1999; DeBoskey & Jiang 2012). However, for constant loss recognition, we expect contrasting over-/underprovisioning effects on current-year non-discretionary LLP, indicated by a negative sign. We find that 68.4 percent of the studies with a loan loss allowance pattern use lagged allowances as a regressor, which is we add ALW_{it-1} to our specification list.[19]

Net loan charge-offs

Beaver and Engel (1996) see current loan net charge-offs as a source of information about future charge-offs (e.g., Cheng et al. 2011). In comparison to non-performing loans, net charge-offs are a less noisy indicator of future losses (e.g., Beck & Narayanamoorty 2013). When net charge-offs are high, current loan quality is rather low and therefore higher provisioning is expected (e.g., Kanagaretnam 2010a). Interestingly, only 43.6 percent of the papers use this pattern group in their EM models. However, since this variable should be a significant driver of the variation in non-discretionary LLP, we add current year net charge-offs, which is the pattern with the highest frequency of use (94.1 percent of the papers using a net charge-off pattern) to our specification list, expecting a positive coefficient.

[19] We note that small differences in the specifications, e.g., including ALW_{it-1} (13 times used) or ALW_{it} (6 times used), might also be due to sample specific considerations of autocorrelation.

Macroeconomic variables

Ahmed et al. (1999) point out that macroeconomic factors might have an effect on non-discretionary LLP. The loan repayment behaviour of companies and private households should be somewhat influenced by the overall economic state of the respective country. In particular, increases in unemployment rates or decreases in GDP may point to a deterioration of the economic situation. In these cases, a decrease in loan quality is highly likely, imposing higher demand for LLP (e.g., Bikker & Metzemakers 2005; Bushman & Williams 2012; Beck & Narayanamoorty 2013). However, Kanagaretnam et al. (2004) propose that inclusion of changes in non-performing loans and net charge-offs will control for macroeconomic effects. What is more, some papers (e.g., Laeven & Majnoni 2003; Kanagaretnam et al. 2014) use year and/or country fixed effects, which work similarly. We regard continuous macroeconomic controls to be more precise and add ΔGDP_{jt} and $\Delta UNEMP_{jt}$ to our specification list.

2.1.2.3 Static LLP models

Based on our specification analysis and discussion, we use the different patterns from the identified patter groups to construct the following first-stage regression framework:

Equation 6

$$\begin{aligned} LLP_{it}/TA_{it-1} = {} & \alpha_0 + \beta_1 NPA_{it-1}/TA_{it-1} + \beta_2 \Delta NPA_{it}/TA_{it-1} + \beta_3 ALW_{it-1}/TA_{it-1} \\ & + \beta_4 TL_{it}/TA_{it-1} + \beta_5 \Delta TL_{it}/TA_{it-1} + \beta_6 NCO_{it}/TA_{it-1} \\ & + \beta_7 \Delta GDP_{jt}/GDP_{jt-1} \\ & + \beta_8 \Delta UNEMP_{jt}/UNEMP_{jt-1} \{+\beta_9 SIZE_{it} + \beta_{10} EBTP_{it}/TA_{it-1} \\ & + \beta_{11} CAPB_{it}\} [+\beta_{12} NPA_{it-2}/TA_{it-1} + \beta_{13} \Delta NPA_{it+1}/TA_{it-1}] + \varepsilon_{it} \end{aligned}$$

where, LLP is loan loss provisions, NPA is non-performing assets, ALW is loan loss allowances, NCO is net loan charge-offs, TL is total loans and TA is total assets for bank i in year t. GDP / $UNEMP$ is gross domestic product / unemployment rate for country j (the respective firm i is located in) in year t. $SIZE$ is defined as the natural logarithm of total assets, $EBTP$ is earnings before taxes and provisioning scaled by total assets and $CAPB$ is a capital adequacy ratio, calculated as tier 1 capital to lagged total assets.

Consequently, we set up four static models using particular parts of the regression framework. Model ***S1*** uses all regressors that are not recorded in brackets ($\beta_1 - \beta_8$). We include the set of further controls for size, pre-provisioning performance and capital ratio in ***S2*** by adding the regressors in curly brackets ($\beta_9 - \beta_{11}$) to the specification. In comparison, ***S3*** includes the additional set of non-performing loans in square brackets, while excluding the set of further controls in curly brackets (include $\beta_{12} - \beta_{13}$, exclude $\beta_9 - \beta_{11}$). At last, model ***S4*** uses all regressors in **Equation 6**.

2.1.2.4 Dynamic LLP models

Models ***S1-S4*** are static regressions. In contrast, some authors (e.g. Laeven & Majnoni 2003; Fonseca & Gonzalez 2008; Bouvatier et al. 2014)[20] introduce dynamic EM models by including the lagged dependant variable (lagged LLP) as an independent variable. This change in the modelling of discretionary LLP assumes autoregressive effects of the first order to have a significant influence on the variation of normal LLP, which have not already been captured by the remaining non-discretionary regressors included in the static LLP models. E.g., Laeven and Majnoni (2003) mention the adjustments in loan loss provisioning when banks approach equilibrium reserve levels.

To investigate the differences in fit between static and dynamic approaches, we define model ***D1-D4*** as dynamic versions of models ***S1-S4***. Therefore, we add β_{14} to each of the specifications of ***S1-S4***:

Equation 7

$$
\begin{aligned}
LLP_{it}/TA_{it-1} = {} & \alpha_0 + \beta_1 NPA_{it-1}/TA_{it-1} + \beta_2 \Delta NPA_{it}/TA_{it-1} + \beta_3 ALW_{it-1}/TA_{it-1} \\
& + \beta_4 TL_{it}/TA_{it-1} + \beta_5 \Delta TL_{it}/TA_{it-1} + \beta_6 NCO_{it}/TA_{it-1} \\
& + \beta_7 \Delta GDP_{jt}/GDP_{jt-1} \\
& + \beta_8 \Delta UNEMP_{jt}/UNEMP_{jt-1} \ \{+\beta_9 SIZE_{it} + \beta_{10} EBTP_{it}/TA_{it-1} \\
& + \beta_{11} CAPB_{it}\} \ [+\beta_{12} NPA_{it-2}/TA_{it-1} + \beta_{13} \Delta NPA_{it+1}/TA_{it-1}] \\
& + \beta_{14} LLP_{it-1}/TA_{it-1} + \varepsilon_{it}
\end{aligned}
$$

[20] Laeven and Majnoni (2003) include a dynamic version (using lag-1 and lag-2 LLP) of their model as an alternative specification.

2.1.2.5 Basic models

Models ***S1-S4*** and ***D1-D4*** represent the state of the art with regard to the estimation of discretionary LLP. Therefore, these specifications should be able to disentangle discretionary LLP from their non-discretionary part to a degree where the resulting proxies are viable for EM propositions. Hence using these models should enable researchers to draw conclusive inferences with significantly low probabilities of type I errors when empirical designs apply sufficient controls.

However, some studies within special settings deliberately leave out non-performing loans (e.g., Cavallo & Majnoni 2002; Bikker & Metzemakers 2005 and Bouvatier et al. 2014, who report data availability as a reason)[21], net loan charge-offs (e.g., Beatty & Liao 2011; DeBoskey & Jiang 2012; Cohen et al. 2014) and loan loss allowances (e.g., Cheng et al. 2011; Bushman & Williams 2012; Bouvatier et al. 2014) as regressors, while numerous studies do so uncommented.[22] To investigate the validity of simplified approaches, we set up basic models. Based on our specification analysis as well as our own data availability, we identify three variables, which are predominantly accountable for the reduction of data sets, namely non-performing loans/assets, allowance for loan losses and net charge-offs. The models are characterised as follows:

Equation 8

$$LLP_{it}/TA_{it-1} = \alpha_0\{+\beta_1\, NPA_{it-1}/TA_{it-1} + \beta_2\, \Delta NPA_{it}/TA_{it-1}\}[+\beta_3\, ALW_{it-1}/TA_{it-1}] + \beta_4\, TL_{it}/TA_{it-1} + \beta_5\, \Delta TL_{it}/TA_{it-1}\,(+\beta_6\, NCO_{it}/TA_{it-1}) + \beta_7\, \Delta GDP_{jt}/GDP_{jt-1} + \beta_8\, \Delta UNEMP_{jt}/UNEMP_{jt-1} + \varepsilon_{it}$$

For model ***B1***, we exclude all brackets (β_{1-3} and β_6) and obtain a model, which is only dependent on the development of two factors: the total loan patterns and macroeconomic effects. For model ***B2-B4***, we include either NPA, ALW or NCO, while for model ***B5-B7*** we include combinations of two out of the three terms. We estimate these basic models to gain insights concerning two interrelated questions. First, which of these basic pattern

[21] Data availability is particularly important for researchers, who focus on bank types other than commercial banks and/or smaller (non-US) samples, due to less pronounced regulations and/or disclosure requirements.
[22] Some studies use the excluded specification parts in untabulated robustness tests (e.g., Bushman & Williams 2012).

groups are important when it comes to splitting discretionary and non-discretionary LLP? Second, does an exclusion of seemingly important variables cause biases and/or lead to substantially varying results on the second stage, especially when it comes to the prediction power of the resulting proxies for EM?

2.1.3 Research design

2.1.3.1 Data and sample selection

We obtain an initial data set containing 14,547 firm-year observations from all 1,445 financial institutions available in Thomson Reuters Datastream with industry codes between 6011-6099 and 6710-6719 for the time-period of 2005-2015. We exclude all observations where annual reports follow any accounting standards but US-GAAP and IFRS, which accounts for a loss of 4,888 observations. We furthermore exclude all observations with insufficient data for the estimation of our models to ensure a constant sample. This step accounts for a loss of 7,076[23], leaving 2,583 firm-year observations as the remaining sample. Since more than 70 percent of the observations are U.S.-banks and regulatory differences between countries, even though having decreased in the post-BASEL I + II era (e.g., Beatty & Liao 2014), may influence the measurement of EM, we decide to use a final sample of 1,854 observations containing only banks from the U.S.[24] We winsorize all incorporated variables at the 1st and 99th percentile to control for outliers that could majorly influence our results. **Table 2** gives an overview of the sample selection procedure and the distribution of observations over the sample period.[25]

[23] In particular, the data availability concerning non-performing loan, net loan charge-off and loan loss reserve patterns leads to a significant sample decrease by about three quarters.
[24] The remaining quarter of the remaining sample comprises 50 countries with very few observations each. In addition, the setting of most parts of the related literature focuses on US banks, too.
[25] We apply lagged and forward variables, which results in a final sample period from 2007-2014.

Table 2 – Sample selection procedure and distribution by years

Selection mode	Number of observations
All companies from 2005-2015	14,547
Less:	
Not US-GAAP or IFRS as reporting standards	4,888
Missing data for estimation of models	7,076
Non-U.S. firm-year observations	729
Final Sample	1,854
Distribution by years	
2007	173 (9.3%)
2008	175 (9.4%)
2009	180 (9.7%)
2010	187 (10.1%)
2011	158 (8.5%)
2012	364 (19.6%)
2013	334 (18.0%)
2014	283 (15.3%)

2.1.3.2 Empirical test procedures

2.1.3.2.1 Empirical design

We test the validity of the discretionary LLP proxies in several analyses. First, we compare the coefficients and the goodness-of-fit of the first-stage regressions. Second, we check for measurement errors of the discretionary LLP measures by applying a method proposed by Dechow et al. (1995) and applied by Peasnell et al. (2000). The test evaluates the rate of rejection of the null hypothesis of no EM when actually no EM exists (type I error). Consequently, we study the rate of incorrect rejection of the null hypothesis. We therefore use a randomly selected sample of 25 percent of the firm-year observations from the total sample[26] and compute a PART variable that takes up the value of one if the firm-year observation is located in the randomly selected sample, and zero otherwise. We study significances of the randomised PART variable, which should not have any significant influence on the dependent variable. We repeat the test procedure 10,000 times for each of the models to cancel out the problem of randomly selecting firm-year observations that actually have significant influences on the dependent variable.

[26] We follow Peasnell et al. (2000) and note that this percentage should capture a typical, dichotomous variable of interest with no systematic influence on our EM proxy. We point to the selection process and the probabilities to draw observations that bear a certain communal characteristic that could lead to a significance, which decreases with increasing number of observations drawn from the sample for our PART variable. We coincide with Peasnell et al. (2000) and Dechow et al. (1995) that this is simply a test of whether the Gaussian assumptions underlying our regressions in **Equation 6**, **Equation 7** and **Equation 8** are satisfied.

Third, we address omitted variable problems of the first stage using two different approaches. First, we study the performance error bias of our coefficients and residuals (e.g., Kothari et al. 2005). Therefore, we use a systematic sampling method (e.g., Dechow et al. 1995) selecting extreme performance firm-year observations and compute another PART dummy variable. We then estimate second-stage univariate regressions once again. When financial performance is one major omitted variable that can majorly improve specification of the model, the coefficients on the PART variable should indicate a significant influence at conventional levels. Peasnell et al. (2000) note that having a systematically selected sample, endogeneity might actually lead to correct significances and therefore false rejections of the models. In particular, observations in the extreme performance parts of the distributions might represent very successful or unsuccessful banks and lead to systematic EM approaches, e.g., very successful banks in the positive extreme part of the distribution might use considerable negative discretionary LLP to build up loan loss allowances and vice versa.[27] We address this problem in more detail in section 2.1.4.3.

We additionally use a more general OMV test proposed by Young (1999). Therefore, we estimate a multivariate second-stage regression containing several omitted and possibly correlated variables.[28] When applying this approach, we must include several conventional control variables to avoid significant coefficients on the applied omitted variables while this variation can actually be explained by the conventional EM control variables.

Finally, we apply a test procedure suggested by Beatty and Liao (2014), who address the prediction power as another desirable property of functional EM proxies. If our discretionary LLP proxies are well fitted, they should be highly correlated with events of actual use of EM. Consequently, if a proxy is able to predict EM events more reliably than another proxy, it can be considered as superior. Due to the lacking observability of actual EM, we employ a proxy for suspected or detected, and therefore highly likely EM. Since 2005, the SEC publicly releases correspondence between SEC staff and SEC filers that

[27] Peasnell et al. (2000) alternatively use decile-specific first-stage regressions and compare abnormal accruals for the extreme decile group with the remaining groups. While recognizing the problem with the test for extreme performance, this alternative test also lacks informative value, since assessment of abnormal accruals from the decile-specific regressions is hard without an unbiased benchmark.
[28] What is important, we assume that these variables could additionally help to separate non-discretionary and discretionary accrual parts.

can be retrieved through the EDGAR database. In these comment letters, the SEC staff comments, criticises and requests more or altered information regarding disclosures.

Therefore, we estimate a univariate logit regression of a dummy variable capturing SEC comment letters and/or restatements on the respective discretionary LLP.

Altogether, we exercise the following steps for each of the presented models:

(1) We estimate first-stage regressions of:

   a. The models ***B1-B7***, models ***S1-S4*** and models ***D1-D4*** using OLS with two-way clustering at the bank and year level (Petersen 2008; Gow et al. 2010).

   b. The models ***D1-D4*** using system GMM with Windmeijer correction (Windmeijer 2005).

   We compare coefficients and apply goodness-of-fit tests for all models.

(2) We compute signed ($DLLP_{it}$) and absolute discretionary LLP ($DLLP\_abs_{it}$) for bank i in year t as the residual and the absolute value of the residual from each model for both estimation procedures.

(3) We construct the following samples for each of the discretionary LLP proxies:

   a. We randomly select 25% of the firm-year observations of the total sample and construct an indicator ($PART_{it}$) that takes up the value of one if the observation has been selected, and zero otherwise.

   b. We select 1%, 5% or 10% of the firm-year observations of the total sample from the firm-years with extreme cash-flow performance and construct an indicator value ($PART_{it}$) that takes up the value of one if the observation has been selected; and zero otherwise.

(4) We estimate the following univariate regression for all models, with PART being defined as stated in (3)a. and (3)b.:

Equation 9

$$DLLP\_abs_{it}(DLLP_{it}) = \alpha + \beta PART_{it} + \varepsilon_i$$

Using OLS regressions with specially designed robust standard errors for heteroscedastic cases[29] and test whether coefficient β is significantly different from zero at conventional levels (0.10, 0.05 and 0.01 levels).

(5) We apply 10,000 iterations for steps (3)a. & (4).

(6) We estimate the following multivariate regression of signed ($DLLP_{it}$) and absolute discretionary LLP ($DLLP\_abs_{it}$) for all models using two-way clustering at the firm and year level (Petersen 2008; Gow et al. 2010):

Equation 10

$$\begin{aligned} DLLP\_abs_{it}(DLLP_{it}) &= \alpha + \beta_1 LLP_{it-1}/TA_{it-1} + \beta_2 SIZE_{it} + \beta_3 EBTP_{it} + \beta_4 CAPB_{it} \\ &+ \beta_5 \Delta GDP_{jt}/GDP_{jt-1} + \beta_6 \Delta UNEMP_{jt}/UNEMP_{jt-1} \\ &+ \beta_7 LLP_{it}/TA_{it-1} + \beta_8 GROWTH_{it} + \beta_9 LOSS_{it} \\ &+ \beta_{10} LOANINT_{it} + \beta_{11} INCDIV_{it} + \beta_{12} CFO_{it-1}/TA_{it-1} \\ &+ \beta_{13} CFO_{it}/TA_{it-1} + \beta_{14} CFO_{it+1}/TA_{it-1} + \varepsilon_i \end{aligned}$$

(7) We estimate the following univariate regression of the comment/restatement of K-10/Q-10 indicator ($CRKQ_{it}$) on the average absolute discretionary LLP ($ARES_{it}$) for all models using a logit regression with two-way clustering at the bank and year level (Petersen 2008; Gow et al. 2010):

Equation 11

$$CRKQ_{it} = \alpha + \beta ARES_{it} + \varepsilon_i$$

And calculate probabilities based on the coefficient β.

The randomised PART variable should not represent any systematic EM, particularly when univariate regressions are run at a high frequency rate. We expect rejection rates of the null hypothesis of no influence of the PART variable in test statistics for coefficient β to be (not) significantly different from 10%, 5% respectively 1% when model specifications are (good) poor at a confidence interval of 90%, 95% respectively 99%. For the

[29] We find significant heteroscedasticity in the sample, which is why we apply standard errors using the Davidson and MacKinnon (1993) method, which obtains more conservative results in cases of heteroscedastic models.

extreme performance PART variable, we assume significant coefficients on PART whenever financial performance drives the magnitude of the discretionary accrual component significantly.[30]

Concerning the multivariate omitted variable regressions, we use the approach proposed by Young (1999), modify it when necessary and use the signed and absolute discretionary LLP. Consequently, we include cash flow from operations as a first omitted and possibly correlated variable. Accruals function as accounting-based adjustments of cash flows to obtain an earnings proxy that captures fundamental firm performance more accurately (e.g., Ball & Shivakumar 2006). Dechow and Dichev (2002) show that when accruals quality is high, accruals' variation should capture performance measurement errors of cash flows. Dechow et al. (1995) and McNichols (2002) further show for the non-financial industry models that excluding cash-flow patterns can cause omitted variable problems. We therefore include lagged (CFO_{it-1}), current (CFO_{it}) and forward (CFO_{it+1}) cash flows from operations as separate regressors to study whether these problems also arise in our financial industry models.[31] Young (1999) uses fixed asset intensity to capture the magnitude of the depreciation expense as one major accrual. We correspond with this idea and use banks' counterpart for LLP, e.g., loan intensity ($LOANINT_{it}$), as another omitted variable. The common banking literature (e.g., Stiroh 2004; Stiroh & Rumble 2006; Lippett et al. 2008) further discusses income diversification ($INCDIV_{it}$), e.g., the ratio of non-interest to interest income, as one important proxy for the strategic alignment of the bank. In particular, we regard this omitted variable as a proxy for the significance of interest income and therefore the significance of the loan portfolio for the bank from an earnings perspective. Following Young (1999), we also integrate $GROWTH_{it}$, calculated as the change in sales scaled by lagged total assets. Expanding banks experience growth in assets as well as liabilities. However, if this development is not overall symmetric, growth may influence the level of non-discretionary accruals disproportionately

[30] As discussed earlier, this could indicate low quality models or actual EM in high-performance situations, e.g., models of high quality. We further discuss this issue when we present the results in section 2.1.4.3.
[31] Even though the relationship between cash flows and earnings might be less direct and intuitive within banks, accruals should also be means of providing a more accurate performance measure, hence the inclusion of the operating cash flow variables should be able to describe non-discretionary LLP variation. E.g., consider the case of a bank with a high-risk loan portfolio reporting high operating cash flows in the first year of observation. Here, high LLP will appropriately counter-steer to provide a smoothed earnings number, while operating cash flows might deteriorate in the following years as high-risk loans are charged off. On a final note, LLP might have a rather forward-looking impact; hence, cash flows in year t+2 or t+3 could be appropriate alternative variables to map the mechanism of provisioning accruals.

and therefore growth needs to be accounted for when modelling discretionary accruals. (e.g., Sloan 1996; Young 1999). While industry models account for the change in sales as growth measure, the banking industry models do not, which is why we include $GROWTH_{it}$ to study the necessity of consideration.

The set of control variables used in the regression is as follows. We add lagged and current total LLP to control for still existing measurement errors. $LOSS_{it}$ is an indicator variable that takes on the value of one if the net income before extraordinary items is negative for firm i in year t, and zero otherwise. It captures the differences in EM incentives when earnings are below zero. Once again, we incorporate the controls for macroeconomic effects ΔGDP_{jt} and $\Delta UNEMP_{jt}$.

For our prediction power test, we follow Beatty and Liao (2014) and construct a dummy variable $CRKQ_{it}$, which equals one if:

a. The bank received an SEC comment letter regarding their annual (K-10) or quarterly (Q-10) earnings report due to unappropriated handling of LLP,

   or

b. The bank restated their annual (K-10) or quarterly (Q-10) earnings report due to reasons connected to the handling of LLP,

and zero otherwise.[32] Based on the resulting coefficients of the univariate regressions, we calculate probabilities. If the resulting probability is above 50%, the model assumes actual EM and expects a detection by the SEC, while probabilities below 50% are interpreted as the absence of actual EM.

2.1.3.2.2 Methodology

We estimate first-stage regressions using two different estimation procedures. Consequently, we use the estimated coefficients to calculate respective residuals of the models for each observation to comprehend our EM proxies (abnormal/discretionary LLP). To estimate the seven basic models, four static and four dynamic models, we use OLS with

[32] Consequently, $CRKQ_{it}$ splits the sample into banks with a high probability of EM in the financial year and banks, which do not or considerably less engage in EM.

standard errors based on two-way clustering at the year- and bank-level (e.g., Petersen 2008; Gow et al. 2010). For the four dynamic models, we alternatively follow prior banking literature with dynamic settings and use system GMM (e.g., Arellano & Bover 1995; Blundell & Bond 1998). We apply all available lags as GMM-style instruments for the lagged dependent variable LLP_{it-1}, while all remaining variables are considered as strictly exogenous.[33] Since standard errors are downward biased in a non-asymptotic setting (e.g., $T \rightarrow \infty$ is not given), and significances therefore overconfident, we apply the standard error correction for finite sample panels developed by Windmeijer (2005).

For models ***B1-B7*** as well as models ***S1-S4*** and ***D1-D4*** using OLS with two-way clustering, we report R^2, adjusted R^2 and Bayesian Information Criterion (BIC) as tests for absolute and relative goodness-of-fit. Higher R^2 (respectively adjusted R^2) values indicate a higher explanatory power of the model and therefore a higher proportion of variation of non-discretionary LLP being explained, leading to lower volatility and therefore more conservatism in our discretionary LLP proxy, which can limit type I errors.[34] Lower BIC values indicate a higher predictive power of the model. Schwarz (1978) introduces the BIC as an extension of the Akaike Information Criterion, assigning different weights to the penalties for the inclusion of additional variables based on the natural logarithm of the number of observations for nested and non-nested models. Given our models that are highly nested in each other, BIC allows us to determine the model with the highest likelihood of generating the underlying data (e.g., Raftery 1995).[35] For the system GMM estimator, we alternatively report AR (1), AR (2) and Hansen statistics as goodness-of-fit tests. They should be significant for AR (1), respectively insignificant on a 10% level for AR (2) and Hansen statistics. These thresholds should not be exceeded to obtain valid regressions and interpretable coefficients.

[33] Bouvatier et al. (2014) also use this assumption in their analysis. Any different approach would contradict the OLS assumptions of exogeneity between dependent and independent variables and question appropriateness of OLS for the remaining models.
[34] We note that this is no clear indication of a better model, although it is likely that regressors, which marginally increase the explanation of the variation in the dependent variable, enhance the modelling until a certain degree of overall explanatory power, since it is highly unlikely that high degrees of total LLP are systematically driven by discretion. Nevertheless, we build our conclusions on various additional tests and not only on the R^2 of the first-stage regressions.
[35] Decisively, we do not seek to explain 100% of the variation in total LLP, but the non-discretionary part. This is why we analyse actual prediction rates for situations of highly likely use of EM in section 2.1.4.5.

2.1.4 Results

2.1.4.1 Descriptive statistics

Table 3 contains descriptive statistics for all relevant variables. Banks in our sample on average provision an amount of 0.47 percent of the lagged total assets, while mean non-performing loans are 2 percent of the lagged total assets. The loan portfolios of the banks represent on average 70.61 percent of the beginning of the year total assets, while variation is rather low with an interquartile range of only 18.01 percent. Mean net charge-offs (non-performing loans) of 0.3559 (0.02) together with a standard deviation of 2.2729 (0.0215) imply that few extreme observations drive the mean.[36]

Table 3 – Descriptive statistics of relevant variables

Descriptive statistics of relevant variables					
	Mean	Q_{50}	**Std. Dev.**	Q_{25}	Q_{75}
Basic specification patterns					
LLP_{it}	0.0047	0.0024	0.0071	0.0009	0.0057
NPA_{it-1}	0.0200	0.0136	0.0215	0.0062	0.0254
ΔNPA_{it}	0.0013	0.0000	0.0136	-0.0039	0.0049
ALW_{it-1}	0.0110	0.0097	0.0060	0.0073	0.0132
TL_{it}	0.7061	0.7001	0.1606	0.6126	0.7927
ΔTL_{it}	0.0454	0.0309	0.1079	-0.0069	0.0737
NCO_{it}	0.3559	0.0020	2.2729	0.0005	0.0067
Set of further control variables					
$SIZE_{it}$	14.4496	14.1297	1.5724	13.3686	15.2051
$EBTP_{it}$	0.0133	0.0138	0.0088	0.0094	0.0177
$CAPB_{it}$	0.0986	0.0948	0.0258	0.0837	0.1088
Macroeconomic control variables					
ΔGDP_{jt}	0.0308	0.0378	0.0184	0.0314	0.0411
$\Delta UNEMP_{jt}$	0.0190	-0.0889	0.2187	-0.0976	0.0319
OMV variables					
CFO_{it}	0.0126	0.0136	0.0292	0.0089	0.0181
$Growth_{it}$	0.0005	-0.0005	0.0088	-0.0033	0.0031
$LOSS_{it}$	0.1348	0.0000	0.3416	0.0000	0.0000
$LOANINT_{it}$	0.6608	0.6760	0.1176	0.5990	0.7431
$INCDIV_{it}$	0.2677	0.2141	0.2360	0.1296	0.3291

All variables are defined in Appendix A.

[36] We note that few banks exist, which account for considerable net-charge offs and run into financial distress in the following year, especially throughout the years of the financial crisis. However, our results do not change when we alternatively exclude these observations.

A mean allowance for loan losses of 1.1 percent compared to mean non-performing loans of 2.0 percent implies that on average banks might be in a situation where building up allowances is an issue. **Table 4** and **Table 5** show signed and absolute discretionary LLP from our models. Signed values are not statistically different from zero on the 0.01 level, indicating that none of our models produces EM proxies with systematic upward or downward bias.[37] Mean Absolute discretionary LLP and standard deviations are considerably different for all models, particularly when system GMM is applied for model ***D2*** and ***D4*** compared to corresponding regressions using two-way clustering. Discretionary LLP from models ***D2*** and ***D4*** with two-way clustering account for the lowest mean values and standard deviations, hence the most conservative proxies.

Table 4 – Descriptive statistics of signed value of discretionary loan loss provisions

Descriptive statistics of the signed value of discretionary loan loss provisions (*DLLP*)					
	Mean	Q_{50}	**Std. Dev.**	Q_{25}	Q_{75}
First-stage OLS					
B1	0.0000	-0.0011	0.0061	-0.0032	0.0015
B2	0.0000	-0.0005	0.0054	-0.0026	0.0015
B3	0.0000	-0.0006	0.0058	-0.0030	0.0018
B4	0.0000	-0.0011	0.0061	-0.0032	0.0016
B5	0.0000	-0.0004	0.0053	-0.0025	0.0017
B6	0.0000	-0.0006	0.0058	-0.0030	0.0018
B7	0.0000	-0.0005	0.0054	-0.0026	0.0016
S1	0.0000	-0.0004	0.0053	-0.0025	0.0016
S2	0.0000	-0.0004	0.0052	-0.0026	0.0016
S3	0.0000	-0.0004	0.0052	-0.0024	0.0015
S4	0.0000	-0.0004	0.0051	-0.0025	0.0016
D1	0.0000	-0.0003	0.0050	-0.0020	0.0013
D2	0.0000	-0.0003	0.0049	-0.0021	0.0014
D3	0.0000	-0.0003	0.0049	-0.0020	0.0012
D4	0.0000	-0.0003	0.0048	-0.0021	0.0012
First-stage GMM					
D1	0.0001	0.0000	0.0053	-0.0017	0.0017
D2	0.0000	-0.0008	0.0083	-0.0047	0.0037
D3	0.0002	-0.0001	0.0058	-0.0028	0.0024
D4	0.0001	-0.0008	0.0079	-0.0046	0.0036

DLLP is the signed value of the residual from the respective first-stage regression. All variables are defined in Appendix A.

[37] We apply untabulated univariate tests to address this issue. They confirm this favorable setting for all discretionary proxies.

Table 5 – Descriptive statistics of the absolute value of discretionary loan loss provisions

Descriptive statistics of the absolute value of discretionary loan loss provisions (*DLLP_abs*)					
	Mean	Q_{50}	**Std. Dev.**	Q_{25}	Q_{75}
First-stage OLS					
B1	0.0039	0.0026	0.0047	0.0012	0.0050
B2	0.0034	0.0021	0.0043	0.0010	0.0041
B3	0.0037	0.0024	0.0045	0.0011	0.0048
B4	0.0039	0.0026	0.0047	0.0012	0.0049
B5	0.0033	0.0021	0.0042	0.0010	0.0041
B6	0.0037	0.0024	0.0045	0.0011	0.0047
B7	0.0033	0.0021	0.0042	0.0009	0.0041
S1	0.0033	0.0021	0.0042	0.0010	0.0041
S2	0.0033	0.0022	0.0040	0.0010	0.0040
S3	0.0032	0.0020	0.0041	0.0009	0.0039
S4	0.0032	0.0021	0.0040	0.0009	0.0039
D1	0.0029	0.0017	0.0040	0.0007	0.0034
D2	0.0029	0.0018	0.0039	0.0008	0.0033
D3	0.0028	0.0016	0.0040	0.0007	0.0033
D4	0.0028	0.0017	0.0039	0.0007	0.0032
First-stage GMM					
D1	0.0031	0.0017	0.0043	0.0008	0.0034
D2	0.0059	0.0043	0.0058	0.0022	0.0077
D3	0.0038	0.0026	0.0044	0.0012	0.0048
D4	0.0056	0.0041	0.0055	0.0021	0.0072

DLLP_abs is the absolute value of the residual from the respective first-stage regression. All variables are defined in Appendix A.

Table 6 shows pairwise correlation matrices for the relevant variables. We find no unexpectedly high correlations that would signify multicollinearity issues, which would lead to implausible or noisy parameter estimates (e.g., O'Brien 2007). We also study mean variance inflation factors and find that they are considerably below five for all specifications.[38] Furthermore, we find expected correlations between our non-discretionary variables and LLP. Interestingly, we find a positive and significant correlation between beginning of the year loan loss allowances and LLP, which could stand for a situation of increasing reserves and loss recognition, which results in even higher provisioning.

[38] A mean variance inflation factor of 5.01 for model ***B1*** could indicate minor multicollinearity problems with this model. However, we follow O'Brien (2007), who extensively discusses the issue of multicollinearity and assesses rules of thumb, in particular a VIF of 10 and above. Altogether, multicollinearity values below 10 should not indicate a crucial problem in our analysis.

Table 6 – Correlations of regressors of the basic, static and dynamic models

Correlations of regressors of the basic, static and dynamic models

		1	2	3	4	5	6	7	8
1.	LLP_{it}	1							
2.	NPA_{it-1}	0.3122***	1						
3.	ΔNPA_{it}	0.3227***	-0.3298***	1					
4.	ALW_{it-1}	0.3031***	0.5944***	-0.1937***	1				
5.	TL_{it}	0.0750***	-0.1068***	0.1820***	0.0320	1			
6.	ΔTL_{it}	-0.2041***	-0.3229***	0.1047***	-0.2455***	0.6942***	1		
7.	NCO_{it}	0.0054	-0.0182	0.1210***	-0.0616***	-0.0181	0.0133	1	
8.	$SIZE_{it}$	-0.0155	-0.1026***	-0.0705***	0.0440*	-0.0358	0.1129***	-0.1379***	1
9	$EBTP_{it}$	-0.1578***	-0.2612***	-0.0015	-0.0747***	0.1706***	0.2271***	-0.0699***	0.2703***
10	$CAPB_{it}$	-0.1543***	0.0147	-0.1187***	0.0679***	-0.0345	-0.0603***	-0.0385*	-0.1437***
11	ΔGDP_{it}	-0.3440***	0.0802***	-0.3430***	0.0840***	-0.0398*	0.0861***	-0.1008***	0.0956***
12	$\Delta UNEMP_{it}$	0.4018***	-0.1206***	0.4067***	-0.116***	0.0600***	-0.1012***	0.1400***	-0.1612***

Correlations of regressors of the basic, static and dynamic models (continued)

		9	10	11	12
9.	$EBTP_{it}$	1			
10	$CAPB_{it}$	0.1090***	1		
11.	ΔGDP_{it}	0.1111***	0.0413*	1	
12.	$\Delta UNEMP_{it}$	-0.1107***	-0.0745***	-0.9342***	1

*, **, *** Indicate statistical significance at the 0.10, 0.05 and 0.01 levels, respectively, based on two-tailed tests. All variables are defined in Appendix A.

2.1.4.2 First-stage regression results

Table 7, **Table 8** and **Table 9** show first-stage regressions of **Equation 6**, **Equation 7** and **Equation 8**. We start analysing the basic models, since they entail concise specifications from where on all other models derive. To assess the absolute and relative goodness-of-fit of our models, we use adjusted r-squared and BIC.[39]

Results in **Table 7** indicate for models ***B1-B7*** that all regressors exert a significant influence on total LLP except for NCO_{it} and ΔGDP_{it}, which are only marginally significant in one case each. The signs for the coefficients are as expected, while the coefficient on ALW_{it-1} is positive and significant for models ***B3***, ***B5*** and ***B6***. This result is in line with our univariate findings and the proposition that higher lagged loan loss reserves are associated with higher loss recognition, leading to higher provisioning in the current year. Concerning the absolute goodness-of-fit, the highest increase in adjusted r-squared is achieved by incorporating the non-performing asset pattern group (increase in R^2_{adj} by 0.16), followed by the inclusion of ALW_{it-1} (increase in R^2_{adj} by 0.07), while adding NCO_{it} leaves an insignificant coefficient and hence no increase in explanatory power. What is more, the mutual use of both NPA and ALW patterns (model ***B5***) results only in a slight improvement in R^2_{adj} compared to model ***B2***. Hence, loan loss allowances and non-performing assets seem to explain equal amounts of variation in total LLP.[40] Taking a closer look at the relative goodness-of-fit, the general results prevail, while the application of BIC unfolds more details. Given the guidelines for evidence in Raftery (1995)[41], model ***B4*** is positively inferior in comparison to model ***B1***, while model ***B7*** outperforms model ***B2*** very strongly. This result still leaves the sole inclusion of our NCO pattern with a loss of fit. However, the mutual use of the NCO pattern and the NPA pattern group strongly indicates a better fit compared to the sole use of NPA, which has not been detected by R^2_{adj}. Even more, compared to absolute goodness-of-fit, we find very strong evidence (BIC difference of 50.49) that there is a better fit of model ***B5*** relative to ***B2***,

[39] We use adjusted r-squared instead of r-squared, since it corrects for the degrees of freedom, while r-squared increases with the inclusion of every new regressor.
[40] In particular, both the coefficients on NPA_{it-1} and ALW_{it-1}, decrease when they are mutually included, representing an adjustment of the reserves based on current non-performing assets.
[41] Raftery (1995) proposes absolute differences from 0-2 as weak, 2-6 as positive, 6-10 as strong and >10 as very strong evidence of preferral of a model in goodness-of-fit.

suggesting that even though NPA pattern groups are already in use, the addition of loan loss allowance patterns can enhance the fit of the model. Consequently, while NPA always seem to increase the fit of our models, loan loss allowances might work as an alternative, while net charge-offs only improve our modelling if they are used together with NPA patterns.[42]

Table 7 – First stage multiple regressions for basic models

First stage multiple regressions for basic models *B1-B7*

Dependent variable: total loan loss provisions (LLP)

$$LLP_{it}/TA_{it-1} = \alpha_0\{+\beta_1 NPA_{it-1}/TA_{it-1} + \beta_2 \Delta NPA_{it}/TA_{it-1}\}[+\beta_3 ALW_{it-1}/TA_{it-1}] + \beta_4 TL_{it}/TA_{it-1} + \beta_5 \Delta TL_{it}/TA_{it-1} (+\beta_6 NCO_{it}/TA_{it-1}) + \beta_7 \Delta GDP_{jt}/GDP_{jt-1} + \beta_8 \Delta UNEMP_{jt}/UNEMP_{jt-1} + \varepsilon_{it}$$

	B1	***B2***	***B3***	***B4***	***B5***	***B6***	***B7***
NPA_{it-1}		0.137			0.108		0.139
		(5.65)***			(7.30)***		(5.59)***
ΔNPA_{it}		0.168			0.168		0.171
		(4.18)***			(4.28)***		(4.19)***
ALW_{it-1}			0.351		0.207	0.351	
			(3.06)***		(2.15)**	(3.05)***	
TL_{it}	0.014	0.008	0.008	0.014	0.005	0.008	0.007
	(3.04)***	(2.79)***	(2.96)***	(3.08)***	(2.61)***	(3.01)***	(2.96)***
ΔTL_{it}	-0.026	-0.013	-0.014	-0.025	-0.009	-0.014	-0.012
	(2.51)**	(1.97)**	(2.30)**	(2.52)**	(1.72)*	(2.31)**	(1.98)**
NCO_{it}				-0.000		-0.000	-0.000
				(1.30)		(1.37)	(1.73)*
ΔGDP_{it}	0.069	0.086	0.101	0.072	0.094	0.105	0.093
	(0.61)	(1.52)	(1.22)	(0.64)	(1.86)*	(1.24)	(1.59)
$\Delta UNEMP_{it}$	0.016	0.016	0.021	0.017	0.017	0.021	0.017
	(1.83)*	(3.30)***	(3.16)***	(1.84)*	(3.97)***	(3.15)***	(3.34)***
$cons$	-0.007	-0.006	-0.008	-0.007	-0.006	-0.008	-0.006
	(1.48)	(2.32)**	(2.07)**	(1.49)	(2.61)***	(2.07)**	(2.38)**
R^2	0.25	0.41	0.32	0.25	0.43	0.32	0.42
R^2_{adj}	0.25	0.41	0.32	0.25	0.43	0.32	0.41
BIC	-13,595.62	-14,032.24	-13,772.22	-13,591.55	-14,082.73	-13,768.00	-14,047.38
F	89.02	85.96	91.05	71.37	80.36	76.11	75.63
N	1,854	1,854	1,854	1,854	1,854	1,854	1,854

*, **, *** Indicate statistical significance at the 0.10, 0.05 and 0.01 levels, respectively, based on two-tailed tests. All test statistics are calculated using two-way clustering at the firm and year level (e.g., Petersen 2008; Gow et al. 2010). We apply the BIC with correction for N (e.g., Schwarz 1978) to assess the relative fit of our models and follow Raftery (1995) in assessment. All variables are defined in Appendix A.

Given these results, model ***S1*** in **Table 8** is of major interest, because it includes NCO, ALW and NPA patterns. The significant coefficients for all three pattern groups demonstrate that integrating NCO_{it} can have a considerable effect if used in addition to loan loss allowances and non-performing assets. ALW_{it-1} is again positive and significant for models ***S1***-***S4***. While R^2_{adj} does not imply any significant difference between modelling without NCO (model ***B5***) and model ***S1***, the difference in BIC shows very strong evidence

[42] Furthermore, the use of net charge-off patterns seems to decrease the fit of the model when we use them together with loan loss allowance patterns, as indicated by a positive evidence for the difference in BIC between model ***B3*** and ***B6***.

for a better fit of ***S1***. Model ***S3*** contains further lagged and forward components of non-performing assets. We observe that these exert only marginal increases in explanatory power, while researchers will have to trade-off a further loss of two panel years. However, relative goodness-of-fit testing shows a difference in BIC that is considerable. The further set of control variables ($SIZE_{it}$, $EBTP_{it}$ and $CAPB_{it}$) likewise explains considerable variation in total LLP, while model ***S4*** contains all static regressors and yields an adjusted r-squared of 0.47, which is the highest explanatory power of the static models. The results on the information criterion are in line, with a difference in BIC between model ***S1*** and ***S4*** of 99.83.

Starting with model ***D1***, the conversion to a dynamic approach expands explanatory power to a level of 0.50 with a very strong evidence of a mutual increase of relative fit. Again, the use of controls and/or non-performing loan regressors has slightly positive effects (R^2_{adj} between 0.51 and 0.53). These results are again supported by a look at the relative goodness-of-fit. What is important, ALW_{it-1} turns insignificant for the dynamic versions ***D1***-***D4***, while LLP_{it-1} is highly significant. We remark that the effect of loan loss reserves is mainly driven by last year's addition to the reserve, e.g., LLP_{it-1}.

Altogether, these results show that non-performing assets are by far the most important pattern group when it comes to modelling EM proxies. The effect of NCO seems to be highly dependent on the ex-ante specification, while the omission of this pattern has no influence on the explanatory power and leads to the lowest decrease in relative goodness-of-fit (difference in BIC between model ***S1*** and ***B5*** of 13.72). Additional lag and forward patterns of the NPA pattern group and/or the additional set of controls significantly improve the BIC, while explanatory power only increases by 2%. However, the switch to a dynamic model yield a much higher increase in both explanatory power (6-7% compared to the respective static model) and BIC (up to a difference of BIC of 272.06). Therefore, EM studies should consider the trade-off between more lags/forwards of NPA patterns and the loss in number of observations, while dynamic models should normally entail insignificant losses of observations.

Table 8 – First stage multiple regressions for the static and dynamic models using two-way clustering

First stage multiple regressions for the static models *S1-S4* and dynamic models *D1-D4* using two-way clustering

Dependent variable: total loan loss provisions (LLP)

$$LLP_{it}/TA_{it-1} = \alpha_0 + \beta_1 NPA_{it-1}/TA_{it-1} + \beta_2 \Delta NPA_{it}/TA_{it-1} + \beta_3 ALW_{it-1}/TA_{it-1} + \beta_4 TL_{it}/TA_{it-1} + \beta_5 \Delta TL_{it}/TA_{it-1} + \beta_6 NCO_{it}/TA_{it-1} + \beta_7 \Delta GDP_{jt}/GDP_{jt-1} + \beta_8 \Delta UNEMP_{jt}/UNEMP_{jt-1} \{+\beta_9 SIZE_{it} + \beta_{10} EBTP_{it}/TA_{it-1} + \beta_{11} CAPB_{it}\} [+\beta_{12} NPA_{it-2}/TA_{it-1} + \beta_{13} \Delta NPA_{it+1}/TA_{it-1}] + \beta_{14} LLP_{it-1}/TA_{it-1} + \varepsilon_{it}$$

	S1	***S2***	***S3***	***S4***	***D1***	***D2***	***D3***	***D4***
NPA_{it-1}	0.110	0.110	0.163	0.157	0.072	0.075	0.117	0.115
	(7.50)***	(7.41)***	(8.70)***	(8.02)***	(5.94)***	(5.97)***	(5.66)***	(5.48)***
ΔNPA_{it}	0.171	0.165	0.158	0.153	0.159	0.155	0.150	0.147
	(4.28)***	(4.28)***	(4.21)***	(4.17)***	(3.84)***	(3.85)***	(3.79)***	(3.78)***
ALW_{it-1}	0.204	0.189	0.251	0.236	-0.014	-0.014	0.032	0.031
	(2.16)**	(2.36)**	(2.75)***	(3.09)***	(0.34)	(0.33)	(0.81)	(0.80)
TL_{it}	0.005	0.006	0.004	0.006	0.005	0.006	0.005	0.006
	(2.70)***	(3.01)***	(2.30)**	(2.66)***	(2.77)***	(2.91)***	(2.33)**	(2.54)**
ΔTL_{it}	-0.009	-0.011	-0.008	-0.010	-0.006	-0.008	-0.005	-0.007
	(1.69)*	(2.11)**	(1.60)	(2.02)**	(1.20)	(1.55)	(1.13)	(1.46)
NCO_{it}	-0.000	-0.000	-0.000	-0.000	-0.000	-0.000	-0.000	-0.000
	(1.76)*	(1.50)	(1.87)*	(1.65)*	(2.10)**	(1.82)*	(2.20)**	(1.96)*
ΔGDP_{it}	0.101	0.110	0.076	0.087	0.053	0.063	0.031	0.041
	(1.92)*	(2.21)**	(1.63)	(1.90)*	(1.50)	(2.09)**	(0.89)	(1.37)
$\Delta UNEMP_{it}$	0.018	0.019	0.015	0.016	0.013	0.014	0.010	0.011
	(3.98)***	(4.49)***	(3.49)***	(3.96)***	(3.65)***	(4.70)***	(2.91)***	(3.79)***
$SIZE_{it}$		0.001		0.000		0.000		0.000
		(2.00)**		(1.82)*		(1.92)*		(1.73)*
$EBTP_{it}$		-0.030		-0.041		-0.024		-0.032
		(1.01)		(1.45)		(0.88)		(1.22)
$CAPB_{it}$		-0.024		-0.022		-0.018		-0.017
		(3.49)***		(3.33)***		(4.46)***		(4.37)***
NPA_{it-2}			-0.080	-0.074			-0.061	-0.057
			(5.92)***	(5.57)***			(3.95)***	(3.46)***
ΔNPA_{it+1}			-0.001	-0.001			0.011	0.011
			(0.06)	(0.03)			(0.66)	(0.63)
LLP_{it-1}					0.379	0.358	0.363	0.344
					(7.37)***	(8.03)***	(6.60)***	(6.93)***
$cons$	-0.006	-0.012	-0.005	-0.010	-0.004	-0.009	-0.003	-0.007
	(2.69)***	(2.17)**	(2.56)**	(1.95)*	(2.57)**	(2.14)**	(2.04)**	(1.81)*
R^2	0.43	0.46	0.45	0.47	0.51	0.52	0.52	0.53
R^2_{adj}	0.43	0.45	0.45	0.47	0.50	0.52	0.51	0.53
BIC	-14,096.45	-14,149.36	-14,151.99	-14,196.28	-14,368.51	-14,412.63	-14,411.68	-14,443.18
F	71.76	57.90	60.89	52.80	76.48	64.10	64.44	57.54
N	1,854	1,854	1,854	1,854	1,854	1,854	1,854	1,854

*, **, *** Indicate statistical significance at the 0.10, 0.05 and 0.01 levels, respectively, based on two-tailed tests. All test statistics are calculated using two-way clustering at the firm and year level (e.g., Petersen 2008; Gow et al. 2010). We apply the BIC with correction for N (e.g., Schwarz 1978) to assess the relative fit of our models and follow Raftery (1995) in assessment. All variables are defined in Appendix A.

For the dynamic models ***D1-D4*** with alternative system GMM estimation in **Table 9**, we apply instrument settings that cancel out issues of over-identifying restrictions, which gives us a certain fit of the models. However, standard goodness-of-fit tests (R^2_{adj}, BIC) do not exist. We alternatively compare coefficients of the models, which leaves us with marginal or no significances for the loans and net charge-off pattern groups.

Table 9 – First stage regressions for the dynamic models using system GMM

First stage regressions for the dynamic models using system GMM				
Dependent variable: total loan loss provisions (LLP)				
$LLP_{it}/TA_{it-1} = \alpha_0 + \beta_1 NPA_{it-1}/TA_{it-1} + \beta_2 \Delta NPA_{it}/TA_{it-1} + \beta_3 ALW_{it-1}/TA_{it-1} + \beta_4 TL_{it}/TA_{it-1} + \beta_5 \Delta TL_{it}/TA_{it-1} + \beta_6 NCO_{it}/TA_{it-1} + \beta_7 \Delta GDP_{jt}/GDP_{jt-1} + \beta_8 \Delta UNEMP_{jt}/UNEMP_{jt-1}$ $\{+\beta_9 SIZE_{it} + \beta_{10} EBTP_{it}/TA_{it-1} + \beta_{11} CAPB_{it}\}$ $[+\beta_{12} NPA_{it-2}/TA_{it-1} + \beta_{13} \Delta NPA_{it+1}/TA_{it-1}]$ $+ \beta_{14} LLP_{it-1}/TA_{it-1} + \varepsilon_{it}$				
	D1	**D2**	**D3**	**D4**
NPA_{it-1}	0.150	0.183	0.057	0.167
	(4.93)***	(4.74)***	(1.07)	(2.32)**
ΔNPA_{it}	0.270	0.276	0.175	0.254
	(4.97)***	(3.95)***	(2.51)**	(3.54)***
ALW_{it-1}	-0.227	-0.191	-0.348	-0.210
	(2.81)***	(1.44)	(2.13)**	(0.93)
TL_{it}	0.006	0.015	0.022	0.017
	(0.68)	(1.38)	(1.58)	(1.13)
ΔTL_{it}	-0.005	-0.024	-0.019	-0.026
	(0.62)	(1.96)*	(1.32)	(1.75)*
NCO_{it}	-0.000	0.000	-0.000	0.000
	(0.88)	(0.76)	(0.43)	(0.71)
ΔGDP_{jt}	0.043	0.038	0.111	0.049
	(1.95)*	(0.88)	(1.84)*	(0.80)
$\Delta UNEMP_{jt}$	0.009	0.008	0.017	0.010
	(3.71)***	(1.73)*	(2.38)**	(1.40)
$SIZE_{it}$		0.002		0.002
		(2.28)**		(2.06)**
$EBTP_{it}$		0.082		0.066
		(0.67)		(0.55)
$CAPB_{it}$		-0.234		-0.212
		(1.98)**		(1.72)*
NPA_{it-2}			0.049	-0.003
			(0.74)	(0.04)
ΔNPA_{it+1}			-0.151	-0.029
			(1.44)	(0.27)
LLP_{it-1}	0.416	0.205	0.439	0.234
	(4.68)***	(1.68)*	(3.83)***	(1.43)
$cons$	-0.004	-0.015	-0.014	-0.019
	(0.77)	(0.85)	(1.79)*	(0.92)
Chi^2	474.64	272.84	354.90	343.71
$Ar1$	-4.08	-3.49	-4.14	-2.76
$Ar2$	-0.03	-1.60	0.67	-0.90
$Hansen$	39.66	26.83	33.62	27.91
J	42.00	42.00	42.00	42.00
N_G	430	430	430	430
N	1,854	1,854	1,854	1,854

*, **, *** Indicate statistical significance at the 0.10, 0.05 and 0.01 levels, respectively, based on two-tailed tests. All test statistics are calculated using system-GMM (e.g., Arellano & Bover 1995; and Blundell & Bond 1998) with all available lags as GMM-style instruments for LLP_{it-1} and Windmeijer (2005) correction, while all remaining variables are considered as strictly exogenous. All variables are defined in Appendix A.

As expected, LLP_{it-1} has a highly significant influence on the variation of current year LLP, surprisingly only as long as we do not include the additional set of control variables. Furthermore, the coefficient on ALW_{it-1} is negative and significant in these models (***D1*** & ***D3***). While this result could challenge our univariate results for ALW_{it-1} as well as the static first-stage regressions, we note that in the case of ALW_{it-1} and LLP_{it-1} capturing similar or identical effects, the sum of the coefficients in our dynamic system GMM models ***D1*** and ***D3*** again shows an overall positive effect of past provisioning.

2.1.4.3 Univariate analysis of measurement errors

Table 10 contains results for the univariate regressions of the EM proxies on our randomized PART variable. For each of the models, we report the frequency of significant β_1 coefficients when we alternatively apply confidence intervals of 1%, 5% and 10%. The frequencies are within the thresholds, e.g., not significantly different from the frequencies you would expect at random. We conclude that the Gaussian assumptions are not violated and we have no biases in this regard.

Table 10 – Univariate analysis of measurement errors for random PART

PART – randomly selected indicator				
Dependent variable: signed discretionary loan loss provisions ($DLLP$) $DLLP_{it} = \alpha + \beta PART_{it} + \varepsilon_i$				
	B1	**B2**	**B3**	**B4**
10% level	954	966	1081	1032
5% level	510	501	529	571
1% level	138	122	107	137
	B5	**B6**	**B7**	
10% level	1032	1044	1023	
5% level	541	541	539	
1% level	118	118	120	
	M1	**M2**	**M3**	**M4**
10% level	1021	984	1055	1026
5% level	521	501	530	538
1% level	115	122	124	112
	M5	**M6**	**M7**	**M8**
10% level	1030	1001	1000	1029
5% level	510	533	537	548
1% level	121	123	132	116
	D1	**D2**	**D3**	**D4**
10% level	999	1006	991	1039
5% level	502	537	477	513
1% level	129	113	118	96

* Indicate whether the number of occurrence of significant coefficients on PART for 10,000 replications of the univariate regression is lower than the hypothesised number. All variables are defined in Appendix A.

Table 11 reports the results for the univariate regression, in which the PART variable represents extreme CFO performance. We find significant coefficients, especially on the 10% and 5% levels. These results could indicate that our EM proxies are biased with regard to extreme CFO performance. If this holds, the applied EM models are not able to unambiguously separate non-discretionary performance influences from discretionary LLP. However, as mentioned earlier, this test can be biased due to EM taking place in observations with extreme CFO performance, e.g., over- or under-provisioning. If this holds, the separation works well and the significant coefficient is the result of successful modelling. In particular, when adding a set of further controls, we include $EBTP_{it}$ as performance before provisioning, while the results on this univariate test do not change.[43]

Table 11 – Univariate analysis of measurement errors for performance PART

PART – extreme performance indicator (CFO)				
Dependent variable: signed discretionary loan loss provisions ($DLLP$) $DLLP_{it} = \alpha + \beta PART_{it} + \varepsilon_i$				
	B1	***B2***	***B3***	***B4***
10% selected	(4.07)***	(4.90)***	(3.78)***	(4.06)***
5% selected	(3.05)***	(3.90)***	(2.94)***	(3.13)***
1% selected	(1.44)	(1.69)*	(1.19)	(1.46)
	B5	***B6***	***B7***	
10% selected	(4.29)***	(3.74)***	(4.75)***	
5% selected	(3.65)***	(2.99)***	(3.84)***	
1% selected	(1.54)	(1.20)	(1.72)*	
	S1	***S2***	***S3***	***S4***
10% selected	(4.13)***	(4.70)***	(4.01)***	(4.41)***
5% selected	(3.58)***	(4.02)***	(3.69)***	(3.90)***
1% selected	(1.56)	(1.81)*	(1.56)	(1.60)
	D1	***D2***	***D3***	***D4***
10% selected	(3.68)***	(4.21)***	(3.76)***	(4.36)***
5% selected	(3.39)***	(3.78)***	(3.42)***	(3.79)***
1% selected	(1.74)*	(1.85)*	(1.67)*	(1.73)*
	DLLP_D1	**DLLP_D2**	**DLLP_D3**	**DLLP_D4**
10% selected	(4.32)***	(3.16)***	(4.27)***	(3.33)***
5% selected	(3.59)***	(2.49)**	(3.71)***	(2.61)***
1% selected	(1.71)*	(0.32)	(1.50)	(0.49)

* Indicate whether the number of significant coefficients on PART for the 10,000 replications is lower than the hypothesised number. Variables are defined in Appendix A.

[43] We further study the influence of cash flow performance in our multivariate analysis in section 2.1.4.4.

2.1.4.4 Multivariate analysis of measurement errors and omitted variables

Table 12, **Table 14** and **Table 16** contain multivariate regressions of the signed discretionary LLP on omitted variables for all models.[44] We use the regressions on signed discretionary LLP to study probable non-linearity issues of variables that have been included in the first stage. In **Table 14**, we find significant coefficients for the further set of controls for size, performance and capital requirements.

Table 12 – Multivariate regressions of signed DLLP on OMV for basic models

Multivariate regressions to study omitted correlated variables in basic models *B1-B7*

Dependent variable: signed value of discretionary loan loss provisions ($DLLP$)

$$DLLP_{it} = \alpha + \beta_1 LLP_{it-1} + \beta_2 SIZE_{it} + \beta_3 EBTP_{it} + \beta_4 CAPB_{it} + \beta_5 \Delta UNEMP_{it} + \beta_6 \Delta GDP_{it} + \beta_7 LLP_{it} + \beta_8 GROWTH_{it} + \beta_9 LOSS_{it} + \beta_{10} LOANINT_{it} + \beta_{11} INCDIV_{it} + \beta_{12} CFO_{it-1} + \beta_{13} CFO_{it} + \beta_{14} CFO_{it+1} + \varepsilon_i$$

	B1	*B2*	*B3*	*B4*	*B5*	*B6*	*B7*
LLP_{it-1}	-0.055	-0.127	-0.200	-0.055	-0.179	-0.199	-0.126
	(4.08)***	(8.12)***	(11.25)***	(4.21)***	(10.03)***	(11.43)***	(8.31)***
$SIZE_{it}$	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	(2.41)**	(2.63)***	(1.88)*	(1.94)*	(2.41)**	(1.50)	(2.46)**
$EBTP_{it}$	-0.000	0.021	-0.003	-0.001	0.012	-0.004	0.020
	(0.01)	(1.05)	(0.30)	(0.08)	(0.71)	(0.34)	(0.97)
$CAPB_{it}$	-0.001	0.003	-0.006	-0.001	-0.000	-0.007	0.002
	(0.54)	(0.43)	(1.98)**	(0.73)	(0.02)	(2.11)**	(0.33)
$\Delta UNEMP_{it}$	-0.019	-0.015	-0.019	-0.019	-0.015	-0.019	-0.015
	(17.53)***	(7.67)***	(16.78)***	(17.98)***	(8.19)***	(16.86)***	(7.63)***
ΔGDP_{it}	-0.092	-0.066	-0.090	-0.093	-0.069	-0.090	-0.066
	(7.80)***	(3.11)***	(7.05)***	(8.05)***	(3.46)***	(7.04)***	(3.10)***
LLP_{it}	0.983	0.854	0.991	0.980	0.863	0.988	0.846
	(65.37)***	(55.51)***	(81.39)***	(59.90)***	(40.97)***	(74.88)***	(48.87)***
$GROWTH_{it}$	0.073	0.058	0.061	0.072	0.059	0.060	0.057
	(4.41)***	(5.39)***	(7.10)***	(4.40)***	(6.61)***	(7.34)***	(4.83)***
$LOSS_{it}$	0.000	-0.001	-0.000	0.000	-0.001	-0.000	-0.001
	(0.11)	(3.68)***	(1.65)*	(0.40)	(3.88)***	(1.49)	(3.79)***
$LOANINT_{it}$	-0.011	-0.008	-0.010	-0.011	-0.008	-0.010	-0.008
	(40.08)***	(14.24)***	(19.05)***	(40.86)***	(13.93)***	(19.94)***	(13.84)***
$INCDIV_{it}$	-0.001	-0.001	-0.002	-0.001	-0.002	-0.002	-0.001
	(3.49)***	(2.43)**	(5.44)***	(2.85)***	(3.24)***	(5.07)***	(2.09)**
CFO_{it+1}	-0.006	-0.007	-0.005	-0.006	-0.007	-0.005	-0.007
	(3.22)***	(2.50)**	(4.11)***	(3.20)***	(3.49)***	(3.76)***	(2.64)***
CFO_{it}	0.004	0.004	0.004	0.004	0.004	0.004	0.004
	(2.26)**	(1.06)	(5.97)***	(2.29)**	(1.61)	(5.19)***	(1.12)
CFO_{it-1}	-0.000	-0.002	-0.011	-0.000	-0.007	-0.011	-0.002
	(0.00)	(0.29)	(2.79)***	(0.03)	(1.23)	(2.72)***	(0.30)
$cons$	0.005	-0.000	0.006	0.005	0.001	0.007	0.000
	(4.58)***	(0.15)	(5.11)***	(5.27)***	(0.58)	(6.14)***	(0.08)
R^2	0.96	0.79	0.92	0.96	0.79	0.92	0.78
R^2_{adj}	0.96	0.79	0.92	0.96	0.79	0.92	0.78
F	2,347.89	174.55	1,401.28	2,328.21	179.67	1,385.67	164.30
N	1,851	1,851	1,851	1,851	1,851	1,851	1,851

*, **, *** Indicate statistical significance at the 0.10, 0.05 and 0.01 levels, respectively, based on two-tailed tests. All test statistics are calculated using two-way clustering at the firm and year level (e.g., Petersen 2008; Gow et al. 2010). All variables are defined in Appendix A.

[44] Compared to our first-stage regressions, we lose one observation each for $GROWTH_{it}$, CFO_{it} and CFO_{it-1} due to data insufficiency, which gives us a sample of 1,851 firm-years.

What is important, these show opposing signs compared to the first-stage results in **Table 8**. Assuming that an inclusion of these regressors should show no significances in the second stage whenever there is a linear relationship, we propose a non-linear fit of the data for these control variables.[45]

Table 13 – Multivariate regressions of absolute DLLP on OMV for basic models

Multivariate regressions to study omitted correlated variables in basic models *B1-B7*

Dependent variable: absolute value of discretionary loan loss provisions ($DLLP\_abs$)

$$DLLP\_abs_{it} = \alpha + \beta_1 LLP_{it-1} + \beta_2 SIZE_{it} + \beta_3 EBTP_{it} + \beta_4 CAPB_{it} + \beta_5 \Delta UNEMP_{it} + \beta_6 \Delta GDP_{it} + \beta_7 LLP_{it} + \beta_8 GROWTH_{it} + \beta_9 LOSS_{it} + \beta_{10} LOANINT_{it} + \beta_{11} INCDIV_{it} + \beta_{12} CFO_{it-1} + \beta_{13} CFO_{it} + \beta_{14} CFO_{it+1} + \varepsilon_i$$

	B1	***B2***	***B3***	***B4***	***B5***	***B6***	***B7***
LLP_{it-1}	-0.039	-0.015	-0.056	-0.039	-0.019	-0.057	-0.016
	(2.94)***	(0.89)	(1.34)	(2.92)***	(0.83)	(1.36)	(0.95)
$SIZE_{it}$	-0.000	-0.000	-0.000	-0.000	-0.000	-0.000	-0.000
	(1.24)	(2.07)**	(1.77)*	(1.22)	(2.09)**	(1.80)*	(2.19)**
$EBTP_{it}$	-0.059	-0.046	-0.059	-0.060	-0.045	-0.060	-0.049
	(2.86)***	(2.86)***	(2.55)**	(3.00)***	(2.62)***	(2.66)***	(3.22)***
$CAPB_{it}$	0.007	0.009	0.012	0.007	0.010	0.012	0.009
	(1.72)*	(2.10)**	(2.18)**	(1.79)*	(1.99)**	(2.17)**	(2.21)**
$\Delta UNEMP_{it}$	0.001	-0.002	-0.001	0.001	-0.003	-0.001	-0.002
	(0.33)	(1.21)	(0.48)	(0.31)	(2.23)**	(0.48)	(1.04)
ΔGDP_{it}	0.001	-0.026	-0.024	0.001	-0.042	-0.024	-0.024
	(0.04)	(1.83)*	(0.87)	(0.02)	(2.80)***	(0.86)	(1.41)
LLP_{it}	0.532	0.426	0.482	0.533	0.409	0.483	0.426
	(11.17)***	(15.38)***	(9.33)***	(10.99)***	(16.90)***	(9.20)***	(15.36)***
$GROWTH_{it}$	0.027	0.002	0.013	0.029	0.002	0.015	0.005
	(2.28)**	(0.27)	(1.53)	(2.35)**	(0.26)	(1.62)	(0.47)
$LOSS_{it}$	-0.001	-0.000	-0.000	-0.001	-0.000	-0.001	-0.000
	(0.94)	(0.23)	(0.57)	(0.96)	(0.13)	(0.62)	(0.40)
$LOANINT_{it}$	0.001	0.000	0.000	0.001	0.000	0.000	0.000
	(0.81)	(0.08)	(0.27)	(0.90)	(0.19)	(0.37)	(0.36)
$INCDIV_{it}$	-0.000	0.001	0.000	-0.000	0.001	0.000	0.000
	(0.29)	(0.97)	(0.65)	(0.36)	(0.89)	(0.59)	(0.71)
CFO_{it+1}	-0.003	-0.001	0.001	-0.003	0.001	0.001	-0.001
	(3.87)***	(0.74)	(1.77)*	(3.78)***	(0.95)	(1.53)	(0.70)
CFO_{it}	-0.001	-0.004	-0.003	-0.001	-0.005	-0.003	-0.004
	(0.58)	(2.56)**	(1.23)	(0.57)	(3.06)***	(1.20)	(2.33)**
CFO_{it-1}	0.006	0.008	0.010	0.006	0.010	0.010	0.008
	(2.24)**	(2.33)**	(2.89)***	(2.36)**	(2.41)**	(2.95)***	(2.56)**
$cons$	0.003	0.004	0.004	0.002	0.004	0.004	0.004
	(1.15)	(2.45)**	(1.52)	(1.12)	(2.40)**	(1.51)	(2.55)**
R^2	0.60	0.51	0.52	0.59	0.48	0.52	0.50
R^2_{adj}	0.59	0.50	0.51	0.59	0.47	0.51	0.50
F	59.11	40.65	43.91	58.88	35.97	43.98	39.64
N	1,851	1,851	1,851	1,851	1,851	1,851	1,851

*, **, *** Indicate statistical significance at the 0.10, 0.05 and 0.01 levels, respectively, based on two-tailed tests. All test statistics are calculated using two-way clustering at the firm and year level (e.g., Petersen 2008; Gow et al. 2010). All variables are defined in Appendix A.

[45] We run simple curve estimations to test this proposition. They show for the respective signed discretionary LLP that common non-linear formulations (e.g., squared or cubic) fit the data significantly better than a linear model.

However, results on absolute discretionary LLP (**Table 15**) show no significant coefficients on size and capital requirement, while the coefficients on $EBTP_{it}$ are significant and negative, indicating pro-cyclical bank provisioning behaviour. Therefore, studies following two-step approaches might consider either non-linear modelling or additional control on the second stage, while one-step approaches are only left with the former solution. The outcomes on LLP_{it-1}, ΔGDP_{jt} and $\Delta UNEMP_{jt}$ also indicate non-linearity (for all models in **Table 12 Table 14** and **Table 16**), hence we suggest an equivalent strategy to address this issue.

Table 14 – Multivariate regressions of signed DLLP on OMV for static and dynamic two-way models

Multivariate regressions to study omitted correlated variables in static models *S1-S4* and dynamic models *D1-D4*

Dependent variable: signed value of discretionary loan loss provisions ($DLLP$)

$$DLLP_{it} = \alpha + \beta_1 LLP_{it-1} + \beta_2 SIZE_{it} + \beta_3 EBTP_{it} + \beta_4 CAPB_{it} + \beta_5 \Delta UNEMP_{it} + \beta_6 \Delta GDP_{it} + \beta_7 LLP_{it} + \beta_8 GROWTH_{it} + \beta_9 LOSS_{it} + \beta_{10} LOANINT_{it} + \beta_{11} INCDIV_{it} + \beta_{12} CFO_{it-1} + \beta_{13} CFO_{it} + \beta_{14} CFO_{it+1} + \varepsilon_i$$

	S1	***S2***	***S3***	***S4***	***D1***	***D2***	***D3***	***D4***
LLP_{it-1}	-0.178	-0.179	-0.182	-0.181	-0.398	-0.386	-0.389	-0.378
	(10.06)***	(10.26)***	(9.74)***	(10.22)***	(27.15)***	(28.07)***	(29.87)***	(31.81)***
$SIZE_{it}$	0.000	-0.000	0.000	-0.000	0.000	-0.000	0.000	-0.000
	(2.27)**	(2.51)**	(1.53)	(2.26)**	(1.59)	(3.17)***	(1.01)	(2.82)***
$EBTP_{it}$	0.012	0.042	-0.003	0.039	0.003	0.027	-0.008	0.025
	(0.65)	(2.34)**	(0.21)	(2.51)**	(0.27)	(2.44)**	(0.74)	(2.31)**
$CAPB_{it}$	-0.001	0.023	0.001	0.023	0.004	0.022	0.005	0.022
	(0.10)	(3.58)***	(0.12)	(3.59)***	(0.97)	(5.27)***	(1.19)	(5.09)***
$\Delta UNEMP_{it}$	-0.015	-0.016	-0.015	-0.016	-0.013	-0.015	-0.013	-0.014
	(8.09)***	(8.74)***	(7.48)***	(8.12)***	(10.22)***	(11.08)***	(9.75)***	(10.57)***
ΔGDP_{it}	-0.070	-0.082	-0.067	-0.078	-0.045	-0.056	-0.044	-0.055
	(3.41)***	(4.01)***	(2.96)***	(3.51)***	(3.07)***	(3.86)***	(2.78)***	(3.50)***
LLP_{it}	0.855	0.858	0.831	0.837	0.873	0.874	0.853	0.856
	(38.24)***	(40.68)***	(48.44)***	(50.76)***	(44.24)***	(46.90)***	(50.90)***	(53.50)***
$GROWTH_{it}$	0.058	0.062	0.059	0.061	0.017	0.022	0.020	0.024
	(5.93)***	(6.68)***	(6.65)***	(7.40)***	(1.52)	(2.14)**	(1.93)*	(2.46)**
$LOSS_{it}$	-0.001	-0.001	-0.001	-0.001	-0.000	-0.000	-0.000	-0.001
	(4.05)***	(3.93)***	(3.52)***	(3.51)***	(2.50)**	(2.63)***	(1.96)*	(2.07)**
$LOANINT_{it}$	-0.008	-0.009	-0.008	-0.009	-0.007	-0.007	-0.006	-0.007
	(13.83)***	(15.81)***	(12.32)***	(14.45)***	(12.62)***	(14.99)***	(10.23)***	(12.29)***
$INCDIV_{it}$	-0.002	-0.002	-0.001	-0.001	-0.000	-0.000	-0.000	-0.000
	(2.93)***	(2.98)***	(3.15)***	(3.19)***	(0.91)	(1.17)	(0.24)	(0.62)
CFO_{it+1}	-0.007	-0.007	-0.006	-0.006	-0.004	-0.004	-0.003	-0.004
	(3.72)***	(3.58)***	(2.90)***	(2.93)***	(2.98)***	(2.92)***	(1.85)*	(1.98)**
CFO_{it}	0.004	0.005	0.002	0.002	0.003	0.003	0.001	0.001
	(1.71)*	(1.72)*	(0.70)	(0.83)	(1.19)	(1.26)	(0.35)	(0.50)
CFO_{it-1}	-0.007	-0.006	-0.006	-0.005	0.000	0.000	0.001	0.001
	(1.18)	(1.12)	(0.86)	(0.86)	(0.12)	(0.06)	(0.28)	(0.21)
$cons$	0.002	0.007	0.002	0.007	0.002	0.006	0.002	0.006
	(0.88)	(3.40)***	(0.96)	(3.00)***	(1.13)	(4.05)***	(1.17)	(3.59)***
R^2	0.78	0.78	0.76	0.76	0.87	0.86	0.84	0.85
R^2_{adj}	0.78	0.78	0.75	0.76	0.86	0.86	0.84	0.84
F	169.70	185.44	150.12	169.08	231.17	237.52	191.55	207.07
N	1,851	1,851	1,851	1,851	1,851	1,851	1,851	1,851

*, **, *** Indicate statistical significance at the 0.10, 0.05 and 0.01 levels, respectively, based on two-tailed tests. All test statistics are calculated using two-way clustering at the firm and year level (e.g., Petersen 2008; Gow et al. 2010). All variables are defined in Appendix A.

Table 13, **Table 15** and **Table 16** show the results for absolute discretionary LLP on omitted variables. For the set of further controls, we find significant coefficients when these have not been added to the first stage of the regressions. Since $EBTP_{it}$ and $CAPB_{it}$, are likely to capture EM incentives of capital management and earnings smoothing, we cannot distinguish between signs of high or low quality of the respective model.

Table 15 – Multivariate regressions of absolute DLLP on OMV for static and dynamic two-way models

Multivariate regressions to study omitted correlated variables in static models *S1-S4* and dynamic models *D1-D4*

Dependent variable: absolute value of discretionary loan loss provisions ($DLLP\_abs$)

$$DLLP\_abs_{it} = \alpha + \beta_1 LLP_{it-1} + \beta_2 SIZE_{it} + \beta_3 EBTP_{it} + \beta_4 CAPB_{it} + \beta_5 \Delta UNEMP_{it} + \beta_6 \Delta GDP_{it} + \beta_7 LLP_{it} + \beta_8 GROWTH_{it} + \beta_9 LOSS_{it} + \beta_{10} LOANINT_{it} + \beta_{11} INCDIV_{it} + \beta_{12} CFO_{it-1} + \beta_{13} CFO_{it} + \beta_{14} CFO_{it+1} + \varepsilon_i$$

	S1	***S2***	***S3***	***S4***	***D1***	***D2***	***D3***	***D4***
LLP_{it-1}	-0.020	-0.028	-0.030	-0.035	0.053	0.045	0.033	0.027
	(0.86)	(1.25)	(1.43)	(1.95)*	(1.05)	(0.96)	(0.82)	(0.70)
$SIZE_{it}$	-0.000	-0.000	-0.000	-0.000	-0.000	-0.000	-0.000	-0.000
	(2.30)**	(1.29)	(1.77)*	(0.84)	(2.45)**	(1.87)*	(2.23)**	(1.49)
$EBTP_{it}$	-0.048	-0.038	-0.046	-0.036	-0.054	-0.042	-0.051	-0.041
	(2.98)***	(3.05)***	(2.88)***	(2.93)***	(3.68)***	(3.20)***	(3.38)***	(3.09)***
$CAPB_{it}$	0.010	0.008	0.011	0.009	0.007	0.006	0.008	0.007
	(2.03)**	(1.13)	(1.97)**	(1.15)	(1.27)	(0.84)	(1.46)	(0.98)
$\Delta UNEMP_{it}$	-0.003	-0.003	-0.003	-0.003	-0.002	-0.003	-0.002	-0.002
	(1.99)**	(2.08)**	(1.88)*	(1.94)*	(1.53)	(1.77)*	(1.49)	(1.65)*
ΔGDP_{it}	-0.040	-0.038	-0.035	-0.033	-0.037	-0.039	-0.032	-0.032
	(2.31)**	(2.20)**	(2.21)**	(2.11)**	(1.87)*	(1.99)**	(1.79)*	(1.87)*
LLP_{it}	0.410	0.400	0.407	0.398	0.359	0.353	0.364	0.358
	(16.82)***	(16.14)***	(14.67)***	(15.21)***	(11.74)***	(12.43)***	(10.12)***	(10.68)***
$GROWTH_{it}$	0.005	-0.003	0.003	-0.003	0.014	0.007	0.009	0.004
	(0.53)	(0.30)	(0.32)	(0.30)	(2.68)***	(0.95)	(1.35)	(0.47)
$LOSS_{it}$	-0.000	-0.000	0.000	0.000	-0.000	-0.000	-0.000	0.000
	(0.29)	(0.06)	(0.03)	(0.14)	(0.42)	(0.16)	(0.04)	(0.13)
$LOANINT_{it}$	0.000	-0.000	0.001	0.000	-0.000	-0.001	0.000	-0.000
	(0.57)	(0.09)	(1.09)	(0.52)	(0.51)	(0.92)	(0.02)	(0.28)
$INCDIV_{it}$	0.000	0.000	0.001	0.000	-0.000	-0.000	0.000	0.000
	(0.71)	(0.49)	(1.01)	(0.83)	(0.26)	(0.49)	(0.15)	(0.13)
CFO_{it+1}	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001
	(0.76)	(0.62)		(3.10)***	(0.85)	(0.92)	(1.11)	(1.11)
CFO_{it}	-0.005	-0.004	-0.006	-0.006	-0.005	-0.005	-0.005	-0.005
	(2.75)***	(2.20)**	(3.68)***	(3.35)***	(1.69)*	(1.93)*	(1.81)*	(2.12)**
CFO_{it-1}	0.010	0.008	0.007	0.006	0.008	0.007	0.006	0.005
	(2.52)**	(1.99)**	(1.99)**	(1.77)*	(2.02)**	(1.81)*	(1.46)	(1.41)
$cons$	0.004	0.004	0.003	0.003	0.005	0.005	0.004	0.004
	(2.48)**	(1.87)*	(1.77)*	(1.32)	(2.45)**	(2.20)**	(2.03)**	(1.70)*
R^2	0.47	0.46	0.48	0.46	0.49	0.48	0.50	0.48
R^2_{adj}	0.47	0.46	0.47	0.46	0.49	0.48	0.49	0.48
F	35.08	32.29	36.02	32.89	30.13	27.87	30.49	27.98
N	1,851	1,851	1,851	1,851	1,851	1,851	1,851	1,851

*, **, *** Indicate statistical significance at the 0.10, 0.05 and 0.01 levels, respectively, based on two-tailed tests. All test statistics are calculated using two-way clustering at the firm and year level (e.g., Petersen 2008; Gow et al. 2010). All variables are defined in Appendix A.

For models ***S2***, ***S4***, ***D2*** & ***D4*** where we include the further set of controls, significances and therefore influences of size and capital adequacy on the inferences drawn from the model are no longer existent.[46] This could also entail disproportionate exclusion of discretionary variation from the total LLP, which would decrease the quality of the model. What is more, we still find significant regressors for $EBTP_{it}$.[47] Altogether, adding a further set of controls should follow careful considerations of which inferences are to be drawn and possible non-linearity of the variables. When we compare the results for the coefficients on the variables $GROWTH_{it}$, $LOANINT_{it}$, $INCDIV_{it}$ and the CFO variables in **Table 12, Table 14** and **Table 16** contain multivariate regressions of the signed discretionary LLP on omitted variables for all models. We use the regressions on signed discretionary LLP to study probable non-linearity issues of variables that have been included in the first stage. In **Table 14,** we find significant coefficients for the further set of controls for size, performance and capital requirements.

Table 12, **Table 14** & **Table 16**, we find significant influences on the variation of non-absolute discretionary LLP. In contrast, we do not find consistent results for the absolute value of discretionary LLP except for CFO_{it} and CFO_{it-1}. Even though all of these omitted and possibly correlated variables do not seem to significantly influence the variation of $DLLP\_abs_{it}$, they seem to explain variation of discretionary LLP to a certain degree, as indicated by the significances for the regressions using signed values. Therefore, including these variables in the first-stage of the EM modelling process could help explain considerable variation of total LLP. If this variation turns out to be mostly non-discretionary, inferences drawn from the second stage could be consolidated. In particular, mapping the relationship between cash-flows and non-discretionary variation of total LLP by including all three (e.g., McNichols 2002) or certain fractions of CFO into a model could improve the quality of our discretionary LLP proxy.

[46] However, when regressions use system GMM in the first stage (**Table 16**), influences of size, performance and capital adequacy are still pronounced, especially when they are included in the first place. Still left with two possible explanations, these results indicate less proper exclusion of variation caused by these variables on the first stage compared to the remaining models.
[47] We are still left with two alternative explanations, hence the lacking isolation of non-discretionary and discretionary LLP and the recognition of EM through the model

Table 16 – Multivariate regressions of signed and absolute DLLP on OMV for dynamic GMM models

Multivariate regressions to study omitted correlated variables in dynamic models *D1-D4* using system GMM								
	Dependent variable: signed value of discretionary loan loss provisions ($DLLP$)				Dependent variable: absolute value of discretionary loan loss provisions ($DLLP\_abs$)			
$DLLP_{it}\ (DLLP\_abs_{it}) = \alpha + \beta_1 LLP_{it-1} + \beta_2 SIZE_{it} + \beta_3 EBTP_{it} + \beta_4 CAPB_{it} + \beta_5 \Delta UNEMP_{it} + \beta_6 \Delta GDP_{it} + \beta_7 LLP_{it} + \beta_8 GROWTH_{it} + \beta_9 LOSS_{it} + \beta_{10} LOANINT_{it} + \beta_{11} INCDIV_{it} + \beta_{12} CFO_{it-1} + \beta_{13} CFO_{it} + \beta_{14} CFO_{it+1} + \varepsilon_i$								
LLP_{it-1}	-0.379	-0.268	-0.424	-0.290	0.064	-0.063	0.042	-0.062
	(15.58)***	(11.52)***	(19.08)***	(13.72)***	(1.31)	(2.56)**	(0.72)	(2.61)***
$SIZE_{it}$	0.000	-0.001	0.000	-0.001	-0.000	0.001	-0.000	0.000
	(1.84)*	(7.14)***	(1.92)*	(7.81)***	(2.08)**	(2.10)**	(1.44)	(2.09)**
$EBTP_{it}$	0.015	-0.057	0.020	-0.044	-0.051	-0.057	-0.059	-0.054
	(0.77)	(2.21)**	(0.91)	(1.80)*	(3.57)***	(1.67)*	(3.19)***	(1.66)*
$CAPB_{it}$	0.010	0.243	0.007	0.220	0.007	0.098	0.007	0.089
	(1.42)	(28.24)***	(1.31)	(28.04)***	(1.34)	(5.90)***	(1.02)	(5.82)***
$\Delta UNEMP_{it}$	-0.010	-0.011	-0.009	-0.011	-0.004	0.006	-0.003	0.005
	(4.51)***	(4.29)***	(3.05)***	(4.29)***	(3.64)***	(2.05)**	(1.52)	(1.79)*
ΔGDP_{it}	-0.035	-0.041	-0.019	-0.042	-0.047	0.068	-0.030	0.059
	(1.37)	(1.48)	(0.52)	(1.40)	(3.02)***	(2.01)**	(1.26)	(1.79)*
LLP_{it}	0.787	0.772	0.872	0.786	0.341	0.242	0.344	0.248
	(27.25)***	(32.02)***	(42.61)***	(38.12)***	(6.92)***	(5.45)***	(7.68)***	(5.90)***
$GROWTH_{it}$	0.014	0.079	-0.008	0.075	0.007	0.002	0.010	-0.002
	(0.66)	(4.43)***	(0.43)	(4.45)***	(0.72)	(0.10)	(0.91)	(0.13)
$LOSS_{it}$	-0.001	-0.001	-0.001	-0.001	0.000	0.002	-0.000	0.001
	(2.32)**	(2.72)***	(1.85)*	(2.56)**	(0.65)	(1.88)*	(0.57)	(1.78)*
$LOANINT_{it}$	-0.006	-0.014	-0.016	-0.015	-0.000	-0.006	-0.001	-0.005
	(6.67)***	(15.47)***	(19.60)***	(18.59)***	(0.04)	(3.85)***	(1.07)	(3.86)***
$INCDIV_{it}$	0.000	-0.001	0.000	-0.001	0.000	-0.001	0.000	-0.001
	(0.05)	(1.71)*	(0.34)	(1.63)	(0.87)	(0.98)	(0.97)	(1.04)
CFO_{it+1}	-0.005	-0.011	-0.008	-0.011	0.001	-0.005	-0.002	-0.005
	(1.73)*	(2.22)**	(3.08)***	(2.37)**	(0.54)	(1.73)*	(0.78)	(1.91)*
CFO_{it}	0.002	0.006	0.007	0.006	-0.005	-0.002	-0.005	-0.002
	(0.54)	(0.92)	(1.46)	(0.98)	(2.15)**	(0.89)	(1.83)*	(0.91)
CFO_{it-1}	0.006	0.005	0.003	0.005	0.008	0.000	0.008	0.001
	(1.06)	(0.59)	(0.46)	(0.59)	(2.24)**	(0.02)	(4.08)***	(0.11)
$cons$	-0.002	0.006	0.005	0.010	0.005	-0.010	0.006	-0.009
	(0.55)	(1.69)*	(1.84)*	(2.71)***	(2.55)**	(1.90)*	(3.07)***	(1.81)*
R^2	0.62	0.80	0.67	0.80	0.42	0.26	0.35	0.26
R^2_{adj}	0.62	0.80	0.67	0.80	0.41	0.26	0.34	0.26
F	62.72	433.92	116.64	417.30	22.57	18.49	19.26	18.83
N	1,851	1,851	1,851	1,851	1,851	1,851	1,851	1,851

*, **, *** Indicate statistical significance at the 0.10, 0.05 and 0.01 levels, respectively, based on two-tailed tests. All test statistics are calculated using two-way clustering at the firm and year level (e.g., Petersen 2008; Gow et al. 2010). All variables are defined in Appendix A.

2.1.4.5 Test for prediction power

Table 17, **Table 18** and **Table 19** include the results of the prediction power test procedure.[48] To analyse the results, we differentiate between two odds. The first value ($Correct\ if\ CRKQ = 1$) indicates the percentage of observations, in which our regression results, given that probable EM had been criticized by the SEC or restated by the firm, correctly predicts EM (Odds > 0.5; positive prediction). The second value shows

[48] We have to exclude all banks without an entry at the SEC, which reduces our sample to 1,826 firm-year observations.

the respective percentage of observations, in which no EM is predicted (Odds < 0.5; negative prediction), given that the SEC or firm had issued no comment or restatement letters on EM related topics. Therefore, a higher value in both percentages indicates a better prediction power.

Table 17 shows the results for basic models. Model ***B1*** accomplishes to positively predict 20.27% cases of potential EM. Adding more explanatory variables into the regression substantially increases the positive prediction power. The only exception is model ***B4***, which is not surprising, as the sole inclusion of NCO_{it} already has no improving effect on the first stage, as indicated by the differences in adjusted R^2 and BIC we discussed earlier. The highest increase in positive prediction power is achieved by including the non-performing assets pattern (NPA_{it-1} and ΔNPA_{it}), either solely (27.99%) or in combination with ALW_{it-1} (28.73%). However, the combination of NPA and NCO patterns results in the highest positive prediction value (28.98%). Regarding negative predictions, the results only vary by a small margin (86.20% to 85.03%). Overall, ***B7*** seems to produce the best trade-off between positive and negative predictions.

The results for static and dynamic models vary by a large margin. The highest positive prediction values are achieved in models ***S1*** and ***S3*** (29.73%). The addition of controls for size, performance and capital requirement decreases the positive prediction power by a large portion (26.99% in ***S2*** and 28.98% in ***S4***), which might underline that these controls (particularly performance and capital adequacy) capture actual EM to some degree. Given the assumption of actual use of EM in every year, an inclusion of these variables results in the removal of discretionary variation from total LLP in the first stage, which leaves the resulting proxy with a lower probability of verifying EM when it actually occurs.

Table 17 – Prediction tests for basic models

Prediction tests for basic models *B1-B7*							
Dependent variable: Comment/restatement of K-10/Q-10 report ($CRKQ$) $CRKQ_{it} = \alpha + \beta ARES_{it} + \varepsilon_i$							
	B1	***B2***	***B3***	***B4***	***B5***	***B6***	***B7***
$ARES_{it}$	125.730	169.846	145.988	124.027	178.302	145.002	169.782
	(6.72)***	(8.23)***	(7.46)***	(6.65)***	(8.39)***	(7.42)***	(8.17)***
$cons$	-0.727	-0.806	-0.778	-0.719	-0.828	-0.773	-0.803
	(8.43)***	(9.69)***	(9.06)***	(8.37)***	(9.84)***	(9.02)***	(9.65)***
$Correct\ if\ CRKQ\ = 1$	20.27%	27.99%	25.50%	20.27%	28.73%	25.50%	28.98%
$Correct\ if\ CRKQ\ = 0$	86.01%	85.03%	86.20%	86.01%	85.32%	85.71%	85.52%
R_p^2	0.02	0.03	0.02	0.02	0.03	0.02	0.03
chi^2	50.02	75.67	62.06	48.89	79.67	61.33	74.59
N	1,826	1,826	1,826	1,826	1,826	1,826	1,826

*, **, *** Indicate statistical significance at the 0.10, 0.05 and 0.01 levels, respectively, based on two-tailed tests. All test statistics are calculated using logistic regressions. All variables are defined in Appendix A.

Table 18 – Prediction tests for static and dynamic two-way models

Prediction tests for static models *S1-S4* and dynamic models *D1-D4*								
Dependent variable: Comment/restatement of K-10/Q-10 report ($CRKQ$) $CRKQ_{it} = \alpha + \beta ARES_{it} + \varepsilon_i$								
	S1	***S2***	***S3***	***S4***	***D1***	***D2***	***D3***	***D4***
$ARES_{it}$	178.313	166.507	177.870	165.114	155.642	154.683	158.496	155.785
	(8.34)***	(7.53)***	(8.18)***	(7.34)***	(7.14)***	(6.84)***	(7.19)***	(6.85)***
$cons$	-0.826	-0.785	-0.806	-0.765	-0.685	-0.683	-0.682	-0.676
	(9.80)***	(9.11)***	(9.66)***	(8.95)***	(8.81)***	(8.58)***	(8.84)***	(8.57)***
$Correct\ if\ CRKQ\ = 1$	29.73%	26.99%	29.73%	28.98%	24.00%	25.25%	25.00%	25.25%
$Correct\ if\ CRKQ\ = 0$	84.93%	86.20%	85.71%	86.69%	86.89%	87.28%	86.50%	87.77%
R_p^2	0.03	0.03	0.03	0.02	0.02	0.02	0.02	0.02
chi^2	78.73	63.13	75.69	60.02	57.58	52.87	58.41	52.92
N	1,826	1,826	1,826	1,826	1,826	1,826	1,826	1,826

*, **, *** Indicate statistical significance at the 0.10, 0.05 and 0.01 levels, respectively, based on two-tailed tests. All test statistics are calculated using logistic regressions. All variables are defined in Appendix A.

Table 19 – Prediction tests for dynamic GMM models

Prediction tests for dynamic models *D1-D4* using system GMM				
Dependent variable: Comment/restatement of K-10/Q-10 report ($CRKQ$) $CRKQ_{it} = \alpha + \beta ARES_{it} + \varepsilon_i$				
	D1	***D2***	***D3***	***D4***
$ARES_{it}$	153.200	62.480	140.209	68.627
	(7.49)***	(5.37)***	(7.35)***	(5.50)***
$cons$	-0.703	-0.607	-0.776	-0.624
	(9.09)***	(7.33)***	(8.97)***	(7.42)***
$Correct\ if\ CRKQ\ = 1$	27.49%	12.31%	23.64%	14.05%
$Correct\ if\ CRKQ\ = 0$	85.13%	88.65%	86.99%	87.77%
R_p^2	0.03	0.01	0.02	0.01
chi^2	63.78	31.68	60.16	33.13
N	1,826	1,826	1,826	1,826

*, **, *** Indicate statistical significance at the 0.10, 0.05 and 0.01 levels, respectively, based on two-tailed tests. All test statistics are calculated using logistic regressions. All variables are defined in Appendix A.

What is more, the introduction of a dynamic modelling already seems to introduce additional bias. Model ***D1*** achieves the lowest positive prediction power, with only 24.00%, which is even lower than most of the basic models. Likewise all other dynamic models produce less appropriate proxies for means of predicting EM. We present two explanations for this result. First, the introduction of an additional lagged dependent variable as a regressor does not add any value to modelling non-discretionary variation in total LLP, but actually colludes with discretionary LLP in the current year. Second, simple assumption of no endogeneity in the models estimated using pooled OLS with two-way clustering tends to lead to biased coefficients for this variable and therefore biases the resulting discretionary proxies. To cancel out the second reason, we alternatively apply the dynamic models with a system GMM estimator. Here, results vary widely. For model ***D1***, the positive prediction value increases from 24.00% to 27.39%, while it is still considerably lower than the prediction value for model ***S1*** (29.73%). For all other models, values decrease substantially, i.e. for model ***D2*** the positive prediction value more than halves from 25.25% to 12.31%.[49] These results could indicate that that our first explanation for the decrease in positive prediction power for dynamic modelling approaches is valid. However, we cannot cancel out that estimation efficiency of system GMM estimators is lacking in our small sample size or the remaining regressors in the first-stage regressions should also be assumed to be endogenous to produce appropriate results (e.g., Bouvatier et al. 2014).[50]

2.1.5 Conclusion

This study tries to analyse modelling of EM in the banking sector. In particular, we analyse the prior literature and identify common specification patterns, examine them and thereby set up various models that capture commonly used specification pattern parts. Consequently, we apply various established statistical methods to test the validity of the models concerning measurement errors, omitted variable biases and prediction power. What is more, we investigate basic models for settings with data limitations, which are

[49] This result is in line with the problem of including the set of further controls discussed earlier. Regarding the negative prediction power, models ***D1*** and ***D4*** result in the best negative prediction values (88.65% and 87.77%), which is likely to be caused by the decrease in positive prediction power, hence no sign of efficient modelling.
[50] As discussed earlier, such an assumption would question most of the applied specifications in bank accounting research, which assume all regressors to be exogenous.

characterised by the absence of potential important specification pattern groups. Applying these models, we additionally try to gain insights with regard to the influence on explanatory and predictive power.

Our results show that the pattern group of non-performing assets is by far the most important influencing factor of non-discretionary variation and therefore a potentially non-disputable data insufficiency variable. The pattern group loan loss allowances can enhance modelling if solely used, but loses most of its explanatory and predictive power when used in combination with the NPA patterns. NCO patterns seem to be dispensable, as they exert only limited influence on non-discretionary variation in first stage regressions and only if used together with NPA patterns. Dynamic models could be superior to static models and should be applied, as there is no loss in observations, if NPA patterns are already used in first stage regression. However, the efficient application of dynamic modelling has to be examined, e.g., choice of estimators. Regarding measurement errors and extreme performance bias, our results indicate that commonly used EM proxies do not suffer from measurement errors but are correlated with extreme CFO performance. Tests of omitted and possibly correlated variables uncover some interesting insights regarding size, performance and capital requirements. All of these variables seem to influence loan loss provisioning in a non-linear way and should therefore be implemented as such. Moreover, our results indicate that the inclusion of certain variables, e.g., variables for growth, income diversification, loan intensity and cash flows may potentially improve the quality of EM proxies.

The results for the prediction power test are somewhat different from first stage results. For our data limitation models, we see an increase in positive prediction power in accordance with the results of the first stage regressions: NPA is a main driver of positive predictive power. Regarding the full-specified static and dynamic models, we see that introduction of $SIZE_{it}$, $EBTP_{it}$ and $CAPB_{it}$ reduces the positive prediction power, while additional lags and forwards of NPA have no effect. Regarding static and dynamic models, the result differs compared to the first stage regression. Positive prediction power decreases by a large margin. Hence, in research environments, in which the prediction power of EM proxies is relevant, static models are superior.

This study also has some possible limitations. The proposed and tested models as well as the strand of literature focusing on residual proxies could suffer from insufficient separation of non-discretionary and discretionary accrual parts on the first stage, making two-step approaches a biased way of modelling. In particular, when correlated non-discretionary variables explaining variation in the discretionary proxy that actually includes non-discretionary variation are not included on the second stage. However, none of our analyses could be applied in a one-step approach and we try to ex-post check for the validity of the models in this regard. Furthermore, we focus only on one accrual, which cancels out noisy variation in the total accrual amount caused by any other accrual we are not able to map with our non-discretionary regressors. However, we propose future research papers to check for the robustness of their results using both one- and two-step approaches, while using residuals from the first-stage for descriptive reasons. Future research should also focus on integrating our omitted and possibly correlated variables to further check for the robustness of their results and possibly improve their modelling of discretionary LLP. This could help to improve the specifications and therefore the modelling process of EM going forward.

2.2 Earnings management modelling in non-financial industries

For the accounting literature and its focus on non-financial industry companies, the development over time is easily despicable. Healy (1985) is the first author to propose a measurement of discretionary accruals. The author assumes that the non-discretionary part of total accruals equals the mean of total accruals during the observed period, indicating that a change in total accruals is attributable to discretionary behaviour.[51] Jones (1991) resolves this restriction and proposes the first estimation-based model, which is able to determine non-discretionary parts from the coefficients of economic figures that reflect the state of the firm. The author uses change in revenue and property, plant and equipment as regressors. This modelling approach follows the idea that the applied economic variables are largely connected with non-discretionary, while the discretionary part can be captured as the residual from the regression. The modified Jones model (DeFond & Subramanyam 1998) additionally considers possible discretionary revenue changes caused by manipulated changes in receivables, while Dechow and Sloan (1991) propose another model, which relaxes the assumption of constant non-discretionary accruals in their so-called „Industry Model". The authors also use an estimation model, but assume that variations in the determinants of non-discretionary accruals are similar in firms of the same industry. Several other attempts to improve the modelling process exist.[52] McNichols (2002) develops one of the more recent models by combining Jones with the quality of accruals model proposed by Dechow and Dichev (2002). The McNichols model therefore adds regressors for lagged, current and forward operating cash flows to the specification and shows that the enhanced model has a significantly higher explanatory power than the Jones model.

Dechow et al. (1995) analyse the applied models and find that the modified Jones model has the highest explanatory power[53], while correlation with financial performance seems

[51] The model by DeAngelo (1986) is a special case of Healy (1985), in which the author assumes that EM is not present in periods without EM incentives, hence defining the non-discretionary part of accruals as the total accruals in the year before the incentive year. Both models suffer under the assumption of permanence of non-discretionary accruals while ignoring the economic context.
[52] See Dechow et al. (2003) and Dechow et al. (1995) for two major attempts to study and develop the discretionary accruals modelling process for the non-financial industries.
[53] However, the authors do not test the validity of the McNichols model. Since McNichols (2002) shows that this alteration is superior compared to the Jones models, this model is applied whenever data sufficiency is not an issue. Since the analysis in chapter 3 would lose a considerable amount of observations, the McNichols model is only used for robustness tests.

to be an issue for all models. On top of that, Kothari et al. (2005) show that performance control can help decrease the probability of over-detection of EM (type I errors) resulting from a high correlation of firm performance and total accruals that would bias discretionary accruals. Therefore appropriate performance control is applied by adding a regressor for firm performance to the models. Alternatively, Kothari et al. (2005) suggest matching firm-year observations of the year and industry and subtracting the discretionary accrual of the matched firm. However, the matching approach can lead to an over-correction for performance. E.g., when a firm has positive discretionary accruals due to high EM and the matched firm might have comparably high positive discretionary accruals, not due to bad fitting of the EM model, but because both firms have comparable EM incentives. Then performance matching would correct for actual EM and hence distort the results on the null hypothesis of no EM in a way that it accepts the null. Kothari et al. (2005) propose that only the even-related EM component is what is of interest for the researcher. When all other incentives as well as the remaining discretionary accrual component are equal, there is no need to worry about over-correction of performance matching. However, this assumption is very strict and many studies test for general EM behavior in special settings instead of only testing for special events like seasoned equity offerings. Performance matching therefore produces "abnormal" EM that is predestined to detect EM performed to meet special incentives that do not exist in the control group. Within the settings in chapter 3 & 4, the integration of an additional regressor for performance controls for the influence of the earnings position of the respective firm in a better way.

The analyses in chapters 3 & 4 use several EM models for robustness reasons. Both studies focus on accrual-based and real activities-based types of EM. The first uses discretionary components of the total accruals of firm i in year t (ACC_{it}). It reflects activities of changing the methods or estimates in financial statements after the fiscal-year end without altering cash flows (e.g., Subramanyam 1996; DeFond & Subramanyam 1998; Dechow & Dichev 2002; McNichols 2002; Dechow et al. 2003; Kothari et al. 2005). Although being not a direct consequence of financial reporting choices, the amount of REM used relates to the quality of the reported earnings numbers. The proxies for REM used in the

analyses follow Roychowdhury (2006) and Gunny (2010)[54], who introduce models that capture various situations where managers alter business actions of the firm. First, overproduction within the firm is used to decrease total costs per unit[55] and consequently costs of goods sold, therefore increasing operating margin and earnings. Operating cash flows (production costs) are then abnormally low (high). Second, firms may conduct sales manipulation in form of accelerated timing, e.g., price discounts, to increase sales volumes in the current period. This will boost earnings for all positive profit margins following sales manipulation. The decreased profit margins entail abnormally high production costs relative to revenues. Firms may also incorporate altered credit terms, leading to lower cash inflows from sales when interest rates are low (or even zero). This entails lower operating cash flows relative to sales compared to ex-ante conditions. Third, real activities may reduce discretionary expenses, which would reduce earnings in the current period. Expenses such as R&D, advertising and selling, general and administrative (SGA) do not convert into earnings in the same period, which makes them a target for REM. The expectations are abnormally low (high) discretionary expenses (operating cash flow) in years of manipulation.

2.2.1 Earnings management models for the analysis in chapter 3

In chapter 3, we use a dataset from 2007-2014 containing all German publicly listed clients of non-Big4 auditors with a total of 1.052 firm-year observations.[56] To compute discretionary accruals, we use the modified, performance-matched Jones model (e.g., Subramanyam 1996; DeFond & Subramanyam 1998; Dechow et al. 2003; Kothari et al. 2005). We therefore include lagged return on assets as a separate regressor in our first-stage regression to control for performance:

Equation 12

$$ACC_{it}/TA_{it-1} = \alpha_0 + \alpha_1\, 1/TA_{it-1} + \beta_1\,(\Delta REV_{it} - \Delta AR_{it})/TA_{it-1} + \beta_2\, PPE_{it}/TA_{it-1} + \beta_3 ROA_{it-1} + \varepsilon_{it}$$

[54] The models by Gunny (2010) are used for robustness reasons in chapter 3.
[55] Under the assumption that fixed overhead costs per unit decrease more than marginal costs per unit possibly increase.
[56] For further details, see section 3.4.1.

We estimate **Equation 12** using industry-year specific cross-sectional regressions.[57] ACC is earnings before extraordinary items minus cash flow from operations. TA is total assets. ΔREV is change in revenues and ΔAR change in accounts receivable during the year and PPE is net property, plant and equipment of the year.[58] ROA_{it-1} is the lagged return on assets.

To account for REM, we collect abnormal production costs, abnormal operating cash-flows and abnormal discretionary expenses. We follow related literature (e.g., Dechow et al. 1998; Roychowdhury 2006; Cohen et al. 2008; Badertscher 2011; Zang 2012) and estimate normal levels of operating cash flows as a linear function of revenues (REV_{it}) and change in revenues (ΔREV_{it}):[59]

Equation 13

$$CFO_{it}/TA_{it-1} = \alpha_0 + \alpha_1\, 1/TA_{it-1} + \beta_1\, REV_{it}/TA_{it-1} + \beta_2\, \Delta REV_{it}/TA_{it-1} + \varepsilon_{it}$$

We keep discretionary accruals estimation procedures for all REM proxies and collect abnormal production costs as the residual of the following model:

Equation 14

$$PROD_{it}/TA_{it-1} = \alpha_0 + \alpha_1\, 1/TA_{it-1} + \beta_1\, REV_{it}/TA_{it-1} + \beta_2\, \Delta REV_{it}/TA_{it-1} + \beta_3\, \Delta REV_{it-1}/TA_{it-1} + \varepsilon_{it}$$

$PROD_{it}$ is the sum of costs of goods sold and change in inventories. Roychowdhury (2006) therefore combines the regressions in **Equation 15** and **Equation 16** for abnormal changes in inventories and abnormal costs of goods sold:

[57] First-stage regressions using cross-sectional estimation can only use one-digit SIC and year combinations due to the limited number of observations on a two-digit SIC and year level, thereby decreasing efficiency of the estimation. On the other hand, estimation on the one-digit SIC code only allows for an aggregated non-stationarity of the coefficients. Therefore, we alternatively use pooled first-stage regressions with two-way clustering and appropriate fixed effects or a one-step approach. We discuss our results in section 3.6 and 3.7.

[58] We follow Roychowdhury (2006) and use an additional scaled intercept (α_1) as a separate regressor. We therefore allow for cases of non-zero discretionary accruals when all other regressors are zero. Additionally, we control for spurious correlation between independent and dependent variable caused by variation in the scaling variable. However, we keep the unscaled intercept, because there is no theoretical reason for the regression to go through the origin. Another important point is that we rely on the r-squared for our goodness-of-fit test, while the reliability of the r-square is questionable when the intercept is suppressed.

[59] Abnormal operating cash flows can be negative and positive when firms use all possibilities of REM. These effects compensate each other and decrease explanatory power (e.g., Chen et al. 2014).

Equation 15

$$\Delta INV_{it}/TA_{it-1} = \alpha_0 + \alpha_1\, 1/TA_{it-1} + \beta_1\, \Delta REV_{it}/TA_{it-1} + \beta_2\, \Delta REV_{it-1}/TA_{it-1} + \varepsilon_{it}$$

The model assumes firm's change in inventory happens when revenues deviate from their expected value of a random walk. This leads to a divergence of inventories from their expected value, resulting in a built up or liquidation of inventory stocks. Roychowdhury (2006) assumes last year's change in sales to influence inventories as well, due to last year's change in inventories, which likewise influences changes in inventories in the current year.

Equation 16

$$COGS_{it}/TA_{it-1} = \alpha_0 + \alpha_1\, 1/TA_{it-1} + \beta_1\, REV_{it}/TA_{it-1} + \varepsilon_{it}$$

The model for costs of goods sold propose that earnings are a fixed proportion of current year's revenues, while the remaining part is regarded as abnormal costs of goods sold. Therefore, we estimate normal levels of these expenses as a linear function of the revenues (Dechow et al. 1998).

We further obtain abnormal discretionary expenses as the residual of the following regression, assuming a linear relation between normal expenses and last year's revenues:

Equation 17

$$DISEXP_{it}/TA_{it-1} = \alpha_0 + \alpha_1\, 1/TA_{it-1} + \beta_1\, REV_{it-1}/TA_{it-1} + \varepsilon_{it}$$

We calculate $DISEXP_{it}$ as the sum of R&D, advertising and SGA expenses.[60]

Following Cohen et al. (2008), Chi et al. (2011) and Greiner et al. (2017), we use a comprehensive measure of REM by combining the three individual standardized measures in the following way (see Kim et al. 2012): $REM_{it} = CFO\_abn_{it} - PROD\_abn_{it} + DISEXP\_abn_{it}$. Consequently, this standardized proxy for REM has a negative scaling, e.g., negative REM entail income-increasing, positive REM income-decreasing EM.

[60] In cases where R&D expenses are not available while advertising and selling, general and administrative expenses are, R&D expenses are set to zero (e.g., Cohen et al. 2008; Ali & Zhang 2015).

Table 20 contains results for industry-specific regressions. Results in the first column show the cross-sectional performance-matched modified Jones model (DeFond & Subramanyam 1998; Kothari et al. 2005), which is used in the regressions reported in chapter 3. Column 2 comprises coefficients for the cross-sectional adapted Jones model (Dechow et al. 2003), which accounts for the expected change in receivables that is assumed not to be due to EM. The model is performance-matched to achieve consistency with the modified Jones model:

Equation 18

$$ACC_{it}/TA_{it-1} = \alpha_0 + \alpha_1\, 1/TA_{it-1} + \beta_1\,((1+k)\Delta REV_{it} - \Delta AR_{it})/TA_{it-1} + \beta_2\, PPE_{it}/TA_{it-1} + \beta_3 ROA_{it-1} + \varepsilon_{it}$$

The slope parameter k represents the part of change in accounts receivable, which can be explained by the change in revenue. It evolves from the following separate industry-year specific regression:

Equation 19

$$\Delta AR_{it}/TA_{it-1} = \alpha_0 + \alpha_1\, 1/TA_{it-1} + k\,\Delta REV_{it}/TA_{it-1} + \varepsilon_{it}$$

The mean (median) value for the slope parameter/coefficient k is 0.1280 (0.1266).[61] The third column shows results for the industry-specific McNichols (2002) model, which is discussed in more detail in the following paragraph on the AEM proxies for chapter 4.

[61] In contrast to Dechow et al. (2003), positive and negative values of slope parameter *k* are included. Importantly, negative values of *k* are at least possible in cases of decreasing (increasing) working capital management performances, which lead to disproportionate positive (negative) changes of accounts receivable when revenues actually decrease (increase).

Table 20 – First stage regression results for discretionary accrual models (chapter 3)

Mean coefficient estimates of discretionary accrual models based on 44 cross-sectional one-digit SIC-year regressions			
	Modified Jones ACC_{it} Coefficient	**Adapted Jones** ACC_{it} Coefficient	**McNichols** ACC_{it} Coefficient
$1/TA_{it-1}$	433.630	437.014	444.745
$\Delta REV_{it} - \Delta AR_{it}$	0.037		
PPE_{it}	-0.024	-0.028	-0.018
ROA_{it-1}	-0.005	-0.003	
$(1+k)\Delta REV_{it} - \Delta AR_{it}$		0.038	
CFO_{it-1}			0.158
CFO_{it}			-0.644
CFO_{it+1}			0.160
ΔREV_{it}			0.111
$intercept$	-0.042	-0.042	-0.021
$Adj.R^2$	0.17	0.17	0.44
N	1,052	1,052	872

We calculate weighted means of the coefficients and adjusted r-squared based on the number of observations of the respective one-digit SIC-year combination. All variables are defined in Appendix C.

Table 21 shows results for the cross-sectional regressions for the REM proxies. $PROD\_abn_{it}$ contains the residual of the cross-sectional regression in column three using **Equation 14**. Abnormal discretionary expenses (column 2) and abnormal operating cash flows (column 1) are the residuals from cross-sectional regressions of **Equation 17** and **Equation 13**.

Table 21 – First stage regression results for REM models (chapter 3)

Mean coefficient estimates of real activities earnings management models based on 44 cross-sectional one-digit SIC-year regressions			
	Abnormal CFO CFO_{it} Coefficient	**Abnormal DISXP** $DISXP_{it}$ Coefficient	**Abnormal PROD** $PROD_{it}$ Coefficient
$1/TA_{it-1}$	-831.524	1,398.817	-911.538
REV_{it}	0.056		0.715
ΔREV_{it}	0.062		-0.000
REV_{it-1}		0.159	
ΔREV_{it-1}			0.061
$intercept$	0.014	0.154	-0.087
$Adj.R^2$	0.28	0.28	0.75
N	1,052	1,052	1,052

We calculate weighted means of the coefficients and adjusted r-squared based on the number of observations of the respective one-digit SIC-year combination. All variables are defined in Appendix C.

We additionally apply the models proposed by Gunny (2010) to comprehend an alternative proxy for REM. Specifically, estimation of discretionary expenses is split into separate regressions for discretionary R&D and SGA expenses:

Equation 20

$$R\&D_{it}/TA_{it-1} = \alpha_0 + \alpha_1\, 1/TA_{it-1} + \beta_1 Size_{it} + \beta_2 Tobin'sQ_{it} + \beta_3\, INT_{it}/TA_{it-1} + \beta_3\, R\&D_{it-1}/TA_{it-1} + \varepsilon_{it}$$

Equation 21

$$SGA_{it}/TA_{it-1} = \alpha_0 + \alpha_1\, 1/TA_{it-1} + \beta_1 Size_{it} + \beta_2 Tobin'sQ_{it} + \beta_3\, INT_{it}/TA_{it-1} + \beta_4\, \Delta REV_{it}/TA_{it-1} + \beta_5(\Delta REV_{it}/TA_{it-1}) * DD_{it} + \varepsilon_{it}$$

Compared to Roychowdhury (2006), Gunny (2010) uses additional market-based size ($Size_{it}$) as a control in both regressions. Greiner et al. (2017) further comment that this market valuation controls for auditor's unrevealed information that has already been incorporated by the market. Gunny (2010) also implements $Tobin'sQ_{it}$, calculated as sum of market value of equity, preferred stocks, long- and short-term debt divided by lagged total assets as a new regressor to capture the effect of the marginal cost to benefit relation that could majorly influence decisions to invest. In addition, INT_{it} controls for the actual funds available for investments, comprehended as sum of income before extraordinary items, R&D and depreciation.

To estimate abnormal R&D expenses in **Equation 20**, $R\&D_{it-1}$ is further included and captures the R&D position of the firm at the beginning of the fiscal year.[62] What is more, to estimate abnormal SGA expenses in **Equation 21**, Gunny (2010) includes two regressors for the change in revenue, which account for the "stickiness" of costs. Anderson et al. (2003) show that increases in costs associated with increases in revenues are more pronounced than the decreases in costs with decreases in revenues due to adjustment costs to decrease cost levels being higher than maintaining costs levels for unutilized resources,

[62] In contrast to Gunny (2010), we claim that estimating a dynamic model using cross-sectional regressions could be biased, since the inclusion of the lagged dependent variable as a regressor results in considerable endogeneity. Therefore, we alternatively estimate this model using system GMM panel estimator by Arellano and Bover (1995) as well as Blundell and Bond (1998) with Windmeijer correction (Windmeijer 2005) assuming $R\&D_{it-1}$ to be endogenous including all lags as possible instruments.

especially when cost decreases are not expected to be long-term. β_5 captures the additional effect for decreases in revenues (DD is an indicator that is one if there is an decrease from t-1 to t, zero otherwise), while β_4 embodies the initial effect.

To collocate alternative abnormal production costs, Gunny (2010) extends the regression proposed by Roychowdhury (2006) by the two already explained components $Size_{it}$ and $Tobin'sQ_{it}$:

Equation 22

$$PROD_{it}/TA_{it-1} = \alpha_0 + \alpha_1\, 1/TA_{it-1} + \beta_1 Size_{it} + \beta_2 Tobin'sQ_{it} + \beta_3\, REV_{it}/TA_{it-1} + \beta_4\, \Delta REV_{it}/TA_{it-1} + \beta_5\, \Delta REV_{it-1}/TA_{it-1} + \varepsilon_{it}$$

We comprehend our alternative REM proxy in accordance with the already discussed negative scaled calculation from the residuals of **Equation 20**, **Equation 21** and **Equation 22**.[63]

2.2.2 Earnings management models for the analysis in chapter 4

In chapter 4, we alternatively use a dataset with all European ASSET4 firms available in the Thomson Reuters database from 2005 to 2014 with 2.733 firm-year observations.[64] To compute AEM, we alternatively use the McNichols (2002) model, which links the discretionary accruals model by Jones (1991) and the accruals' earnings quality model by Dechow and Dichev (2002):

Equation 23

$$ACC_{it}/TA_{it-1} = \alpha_0 + \alpha_1\, 1/TA_{it-1} + \beta_1\, CFO_{it-1}/TA_{it-1} + \beta_2\, CFO_{it}/TA_{it-1} + \beta_3\, CFO_{it+1}/TA_{it-1} + \beta_4\, \Delta REV_{it}/TA_{it-1} + \beta_5\, PPE_{it}/TA_{it-1} + \beta_6\, \Delta GDP_{jt}/GDP_{jt-1} + \varepsilon_{it}$$

We estimate **Equation 23** cross-sectionally for each industry-year combination based on the two-digit SIC code. ACC_{it} is calculated as earnings before extraordinary items minus the cash flow from operations. CFO_{it} (CFO_{it-1}, CFO_{it+1}) is the net cash flow from operations in year t (t-1, t+1). ΔREV_{it} is the change in revenues, and PPE_{it} is net property,

[63] First-stage regression results are comparable to Gunny (2010) and are not tabulated here.
[64] For further details, see section 4.3.1.

plant and equipment. All these variables are scaled by lagged total assets (TA_{it-1}). Continuous control for macroeconomic effects is used due to the lack of a number of observations for an industry-year-country specific estimation. The study follows Cohen et al. (2008) and applies changes in the gross domestic product (ΔGDP_{jt}) of the respective country j, in which firm i is located, for year t.[65] Consequently, we comprehend our AEM proxy, absolute discretionary accruals ($DACC\_abs_{it}$), as the absolute value of the residual from **Equation 23**. To measure the magnitude of REM (e.g., Roychowdhury 2006; Cohen et al. 2008; Cohen & Zarowin 2010; Badertscher 2011; Zang 2012) we again follow the approach by Roychowdhury (2006).[66]

Table 22 contains results for industry-specific regressions. Results in the first column show the McNichols (2002) model, which we use throughout the analysis in chapter 3. Column 2 and 3 show the coefficients for the cross-sectional modified and adapted Jones models, which are used in robustness analyses.[67]

Table 22 – First stage regression results for discretionary accrual models (chapter 4)

Mean coefficient estimates of discretionary accrual models based on 160 cross-sectional two-digit SIC-year regressions			
	McNichols ACC_{it}	**Modified Jones** ACC_{it}	**Adapted Jones** ACC_{it}
	Coefficient	Coefficient	Coefficient
$1/TA_{it-1}$	2,302.885	-1,596.643	-2,594.696
CFO_{it-1}	0.223		
CFO_{it}	-0.645		
CFO_{it+1}	0.160		
ΔREV_{it}	0.065		
PPE_{it}	-0.031	-0.067	-0.066
ΔGDP_{jt}	0.186	0.145	0.130
$\Delta REV_{it} - \Delta AR_{it}$		0.001	
ROA_{it-1}		0.205	0.209
$(1+k)\Delta REV_{it} - \Delta AR_{it}$			0.006
$intercept$	-0.035	-0.060	-0.059
$Adj.R^2$	0.40	0.19	0.18
N	2,733	2,733	2,733

We calculate weighted means of the coefficients and adjusted r-squared based on the number of observations of the respective one-digit SIC-year combination. All variables are defined in Appendix D.

[65] Various dichotomous country indicators control for equivalent effects but significantly decrease the estimation efficiency due to the overall limitation of observations within each two-digit SIC-year grouping.
[66] See chapter 2.2.1 for a detailed discussion of all other models. To account for the differences in the samples, we additionally include changes in gross domestic product to account for country specific macroeconomic effects.
[67] The mean (median) value for the slope parameter/coefficient k is 0.1561 (0.1464). Again, we include positive and negative values of k.

Table 23 shows results for the cross-sectional regressions for the REM proxies. Again, we report mean coefficients of the cross-sectional regression of **Equation 13** in column one, **Equation 17** in column two and **Equation 14** in column three.

Table 23 – First stage regression results for REM models (chapter 4)

Mean coefficient estimates of real activities earnings management models based on 160 cross-sectional two-digit SIC-year regressions

	Abnormal CFO $\boldsymbol{CFO_{it}}$	**Abnormal DISXP** $\boldsymbol{DISXP_{it}}$	**Abnormal PROD** $\boldsymbol{PROD_{it}}$
	Coefficient	Coefficient	Coefficient
$1/TA_{it-1}$	11,337.080	41,966.920	-43,073.020
REV_{it}	0.042		0.810
$DREV_{it}$	0.110		-0.057
ΔGDP_{jt}	0.015	-0.459	-0.478
REV_{it-1}		0.130	
$DREV_{it-1}$			-0.101
$intercept$	0.075	0.140	-0.210
$Adj.R^2$	0.30	0.31	0.84
N	2,733	2,733	2,733

We calculate weighted means of the coefficients and adjusted r-squared based on the number of observations of the respective one-digit SIC-year combination. All variables are defined in Appendix D.

3 The influence of client importance on accrual and real activities earnings management in the small and midsized audit market for listed clients[68]

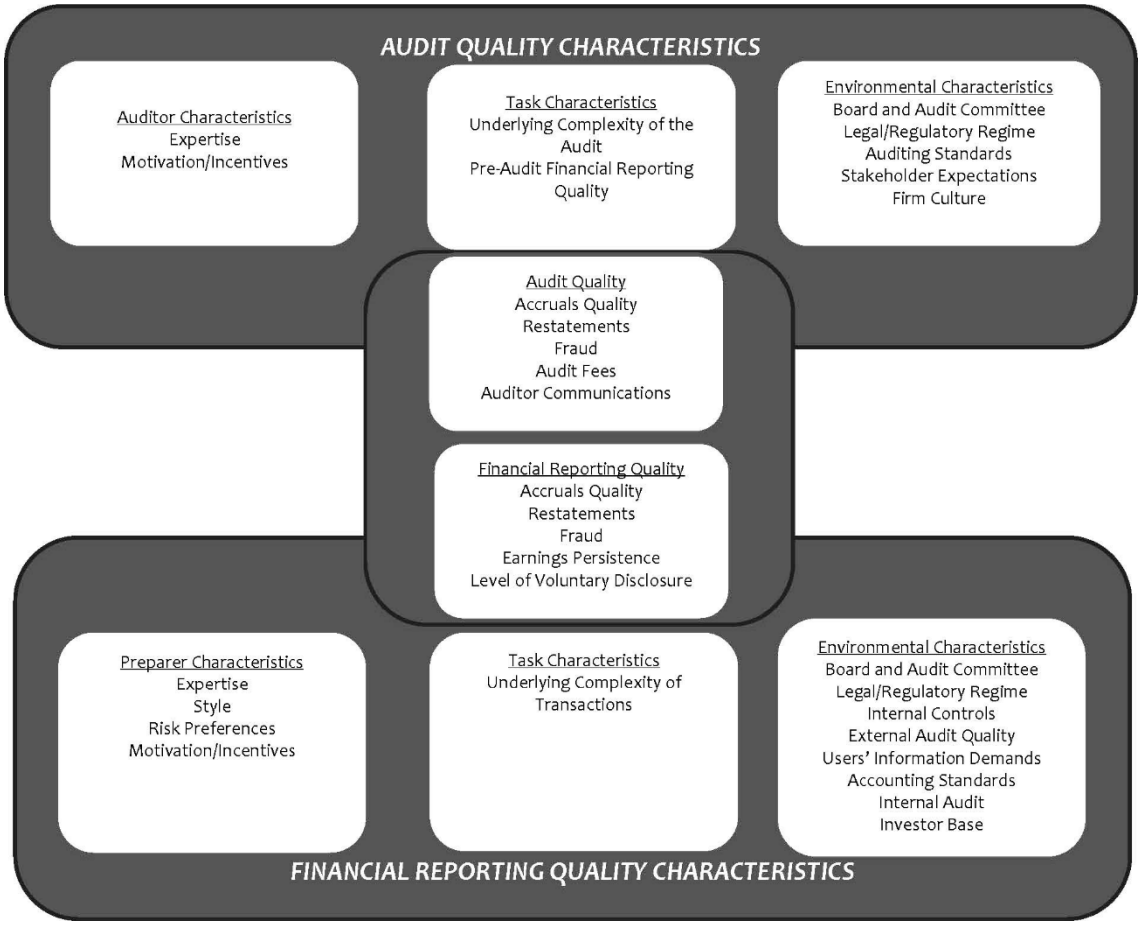
> "Judging whether something is material or not is important, because it determines whether information is included or excluded from the financial statements. […] We know this can be tricky." Hoogervorst (2017)

The IASB aims at providing intensive guidance on how companies should apply standards. However, in the case of opportunistic behaviour by the management, it is inevitable not only to provide thorough standards, but also further ensure reliability of the information disclosed, i.e., whether the decision about the materiality of a transaction is reasonable.

Gaynor et al. (2016) distinguish between preparer's characteristics, task characteristics and environmental characteristics in their framework shown in **Figure 3**. Among others, management's opportunistic incentives can be seen as important preparer characteristics (Dechow et al. 1996; Cheng & Warfield 2005; Bergstresser & Philippon 2006). The complexity of the business transactions of the firm and the attached degree of reporting discretion shape the task characteristics of financial reporting (Dechow & Dichev 2002; Bratten et al. 2016). In addition, various environmental characteristics exist, e.g., internal controls and internal audit, the legal system, the supervisory board and the audit committee. These can already help ensure the conformity of accounting decisions (Dechow et al. 1996; Beasley 1996; Klein 2002; Leuz et al. 2003; Prawitt et al. 2009). However, the ambition to provide reliability of the disclosed information has raised the demand for external auditing. Again, certain auditor and task characteristics exist, while environmental characteristics prevail.

[68] This chapter is based on a 2017 version of a working paper titled "Exerting oneself in the market for PIEs? A fresh look at client importance for small and midsized auditors", which is co-authored by Hansrudi Lenz. The reasoning, results, and interpretations of this might change after the submission of this thesis and the completion of the doctoral degree. The most recent version of this study is available upon request. Please do not cite this working paper without permission.

Figure 3 – Input and environmental variables for audit quality and financial reporting quality



Source: Gaynor et al. (2016), p. 8.

Given the expertise and pursuit of profit of the auditor together with the underlying complexity of the firm and therefore the financial information (Bonner 2008), the auditor pursues to confirm or enhance the quality of the financial reporting provided by the management. To be able to rely on audited financial information, it is important to study audit quality and its influence on the financial reporting quality, e.g., earnings management (EM). The following study deals with the independence of the auditor and question whether we can draw inferences from the degree of dependence on the quality of the audit and therefore the amount of EM possibly used by the management.

3.1 Introduction

Business activities of audit firms create an economic dependence from audit clients. An audit firm intends to keep existing client relations intact over time and looks for new

clients. If an auditor does not agree with management's or the supervisory board's view in accounting subjects and the dispute eventually leads to a modified, e.g., qualified or adverse opinion, she runs into the risk of losing the client because the responsible body of the client may threat the auditor with the loss of audit and non-audit services in the following years (e.g., Leffson 1995; European Commission 1996; Tepalagul & Lin 2015). This threat in conjunction with the intense competition in the audit market (e.g., Köhler et al. 2010; Zülch et al. 2010) may entice the auditor to acquiesce to the client's demands. If the auditor does so she runs into reputation and/or litigation risks, e.g., she will lose future income streams from other audit clients. However, if the revenues from a single audit client represent a material percentage of the total revenues of the audit firm and a compensation of a potential client loss is not possible then there is a material threat with respect to the economic independence of the respective audit firm. Accounting research uses various terms for this phenomenon, e.g., client importance, economic bonding, independence or dependence threat, while most studies regard it as a crucial input factor for the quality of the audit (e.g., Gaynor et al. 2016).

Furthermore, DeAngelo (1981a,b) proposes that small and midsized audit firms, having a smaller number of clients and rather empty pockets (e.g., DeFond & Zhang 2014), are more likely to run into the risk of economic bonding compared to their large Big N counterparts. The risk of losing other audit clients in the case of a following audit failure should be comparably lower than the financial consequences of the loss of a significant single large audit client (e.g., DeAngelo 1981b; DeFond & Zhang 2014). What is more, the financial loss is even higher if the auditor not only loses the audit fees but also the non-audit fees resulting from other permissible assurance and consulting services. Additional non-audit services lead to higher total revenues from a single audit client and are oftentimes more lucrative than audit services, which is also why the economic interests of an audit firm may gradually shift from audit to non-audit services.[69] The threat is that these financial incentives might dominate professional ethical standards of conduct, such as

[69] Some auditors might see the mandatory audit only as a commodity, e.g., an "exchangeable service with the price as the central distinction criteria" (Naumann 2008), resulting in an increase of low balling incentives in the competition for new clients (e.g., prizing the audit below costs; DeAngelo 1981a) in exchange for future quasi-rents, particularly from non-audit services.

objectivity and due care, and culminate in impaired audit quality (e.g., Ludewig 2002; Pfitzer 2006; European Union 2014).

However, prior empirical studies focus on studying financial independence threats based on published audit fees of listed Prime Standard firms (e.g., Bauer 2005; Lenz et al. 2006a; Krauß 2008; Petersen & Zwirner 2008; Zimmermann 2008; Bigus & Zimmermann 2009; Zülch et al. 2009; Zülch et al. 2010; Lopatta et al. 2015), while these PIEs are predominantly audited by Big N auditors (e.g., for Germany the Big4 and BDO). In contrast, there are only few studies with special emphasis on smaller audit firms, partly because these auditors mainly provide audit services to firms that are not listed in the most prominent stock exchange segments. One reason for this focus on Big N auditors is their market share in the important U.S. market, while for other settings data insufficiencies have been prevalent for smaller audit firms.[70] Decisively, since small auditors only have a small number of PIE clients, detailed information about the total revenue of the auditor is inevitable to calculate reasonable client importance indicators for non-Big N audit firms. Most prior studies measure client importance only based on a specific market segment, e.g., they use the sum of total revenues for an audit office or the audit firm from this specific setting as the denominator and do not consider the revenues from all clients of the office or firm.[71]

Consequently, major parts of the related literature find higher quality audits for clients audited by Big N auditors (e.g., Francis et al. 1999; Kim et al. 2003; Farber 2005; Lennox & Pittman 2010), while the European Commission explicitly asked in its Green Book whether the maximum fee from a single client should be restricted (e.g., European Commission 2010). As a result, the European legislator[72] has decided that if total fees received from a public-interest entity (PIE) in each of the last three consecutive financial years are more than 15 % of the total fees received by the statutory auditor in each of these financial

[70] E.g., for Germany the legislator introduced the duty to publish transparency reports according to sect. 55c WPO (old version) with some aggregated financial information only in April 2008.
[71] Notable exceptions are Chung and Kallapur (2003) for the U.S. Big4 market and Sattler (2011), Quick and Sattler (2011a,b) and Lopatta et al. (2015) for the German Big4 market using total revenues from all clients as a correct denominator for client importance ratios. We follow this line of research.
[72] According to Art. 10, par. 3 EU Regulation (EU) No 537/2014 (European Union 2014) for all financial years following June 2016.

years, this fact has to be disclosed and discussed with the audit committee with respect to independence threats and possible safeguards.

Based on this elaboration, we deliberately choose a regulatory-driven threshold approach and study the importance of clients for small and midsized auditors with greater importance of economic bonding. In particular, we apply various regulatory thresholds for excessive client importance and study whether lower audit quality is related to lower financial reporting quality, which we proxy with higher EM.

We aim at describing and assessing the strength of the economic bonds between the auditor and the clients for small and midsized German audit firms based on information from the transparency reports in the years 2009 to 2016 and the respective financial statements of the audited entities. Being able to observe the total sales from all services and all clients for each of the small auditors, we apply a precise proxy for the client importance of small and midsized auditors, e.g., we comprehend CIS as total sales from the client in relation to total sales of the audit firm from all clients and. Furthermore, we apply the commonly used $CINAS$ measured as non-audit sales in relation to total sales from the client. Furthermore, the German has several other advantages for an analysis of small and midsized auditors, e.g., they have a significant market share based on the number of listed clients and there are relatively liberal restrictions with respect to non-audit services, while the audit environment is less litigious.

Consequently, we apply non-linear client importance variables and study the effect on accrual earnings management (AEM). We further allow for another possible strategy and explicitly consider a trade-off between AEM and real earnings management (REM) for cases when AEM is limited/enhanced (Zang 2012). For this purpose, we analyze 1.052 firm-years covering the years from 2007 to 2014 from all German PIEs being audited by a small or midsized auditor at least once in the sample period. The results indicate for CIS that firms with auditor's client importance below the thresholds use lower amounts of AEM and higher amounts of REM as client importance increases. However, for firms with auditor's client importance above 5%, 10% or 15% for one, two or three consecutive years, we find that AEM increases, while the relative weight of REM (trade-off REM vs.

AEM) decreases again. We find that these effects are more pronounced for higher thresholds used for separation as well as for increasing number of consecutive years. However, for $CINAS$ we find no significant results, highlighting the distinction between non-audit service and total client importance proxies. What is more, we provide a reasoning for our results, which particularly signifies the eligibility of PIE engagements by small and midsized auditors and provides a fresh look at client importance for these auditors. Our results hold in various robustness tests and additional analyses, e.g., the inclusion of a best practice audit committee, de-facto office and partner level analyses and when focusing only on extreme EM.

This paper contributes to the literature about the relation between client importance and financial reporting quality in several ways. First, we provide the first empirical evidence that regulatory (audit and non-audit) importance thresholds might embody non-linear audit quality effects. Our data indicates that overstepping certain thresholds of total client importance impairs audit quality and leads to more AEM. What is important, below these thresholds we observe a negative relationship between client importance and AEM. We conjecture that small and midsized audit firms, which successfully obtain a PIE engagement, have higher reputation and litigation incentives[73] that preserve independence up to certain thresholds where dependence excels and leads to higher AEM. We further find that clients seem to adapt their EM strategies by trading-off from AEM to REM when AEM is limited by high audit quality. We further control for effects of the supervisory board or audit committee, which may be a safeguard to impaired auditor independence due to an economic bond between auditor and client.

Second, to the best of our knowledge this is the first paper analyzing the influence of client importance on the trade-off between AEM and REM. Therefore, we are able to answer the question of higher or lower earnings quality thoroughly and raise the awareness and discussion about the possible substitution of AEM by REM when clients experience high audit quality. In addition, in our additional analyses we provide interesting results that best practice audit committees can help mitigate this trade-off issue.

[73] In addition, the auditors know that they are under the spotlight of the German Auditor Oversight Commission and their clients are under the spotlight of the Financial Reporting Enforcement Panel.

Third, we contribute to the discussion of whether small and midsized auditors deliver high or low quality audits, particularly when they pursue PIE audits. This might be important for future discussions and considerations by clients and regulatory bodies to enable more PIE mandates for small and midsized auditors. Decisively, we raise the awareness for the magnitude of non-PIE mandates and the previous overestimation of the economic bond when it comes to discussing PIEs and small and midsized auditors. All of these contributions can be of major interest for standard setters and various stakeholders.

The following chapter briefly discusses the relevant mandatory framework for client importance, the disclosure of audit and non-audit fees in financial statements and the disclosure of information in transparency reports in Germany. Chapter 3.3 discusses the relationship between economic bonding, audit quality and financial reporting quality. We consider the relation between auditor size and economic dependence, present our theoretical framework and discuss methodology and results of related previous studies about client importance as audit quality measure and financial reporting quality, whereinafter we develop our hypotheses. Chapter 3.4 discusses our measures for audit quality, e.g., client importance, and earnings quality, e.g., EM. We characterize the sample, the relevant data and the empirical model. Chapter 3.5 presents descriptive statistics and univariate results, while chapter 3.6 reports our multivariate tests. In chapter 3.7, we provide various additional analyses and robustness tests. Chapter 3.8 summarizes the results and discusses the limitations of our study.

3.2 Client importance – mandatory framework in Europe

To secure the economic independence of auditors, the German legislator has issued specific norms in the German Commercial Code (Handelsgesetzbuch, HGB) and in the German Public Accounting Act (Wirtschaftsprüferordnung, WPO). Additionally, the German Chamber of Auditors (Wirtschaftsprüferkammer, WPK), a public law entity, specifies the professional duties when auditors are carrying out mandatory audits in the Professional Charter for auditors (Berufssatzung WP/vBP).[74] These regulations – based on the Auditor

[74] According to Art. 57 para. 4 WPO these rules are legally binding for auditors.

Directive of the European Union – have been triggered by national and international accounting scandals.[75]

3.2.1 Independence threats resulting from economic bonds

The most relevant norms for securing the financial dependence of auditors in Germany in the period under study are Art. 319 para. 3 sent. 1 no. 5, para. 4 and 5 HGB and Art. 319a para. 2 no. 1, para. 2 HGB. If an auditor has earned in the last five years more than 30 % of his total revenues from professional services from a single audited entity or from entities the audited entity holds more than 20 % of the shares, and it is expected that the auditor also earns more than 30 % in the current year, then the auditor is not allowed to act as an auditor (cf. Art. 319 para. 2 no. 5). This stipulation holds for a sole practitioner, partnerships or corporations of professional accountants and is applicable for the audit of single statements and consolidated financial statements but is not applicable to network firms of the auditor (cf. Art. 319 para. 4, 5 HGB).[76] If the audited entity is a capital market oriented entity according to Art. 264d HGB, then the critical threshold reduces from 30 % to 15 % of the total revenues because of the more pronounced public interest in such firms.[77] An undercut in a single year within the timespan of six years is sufficient to escape from this rule. It is notable that the German government, in its reasoning for this legal rule, has admitted that the 15 % limit is an upper limit with respect to internationally discussed bandwidths from 5 % to 15 %.[78] The argument in favor of this moderate threshold is that smaller audit firms should also have the chance to win larger audit engagements, whereas for large audit firms obeying even a stronger limit should be no problem.[79] These current German legal rules (15 % threshold over six years) with respect to client dependence of auditors for listed clients could be criticized because first, the percentage is very high and second, a time period of six years seems to be too long.

For financial years after June 2016, the above-mentioned 15 % threshold has been replaced by Art. 10 par. 3 Regulation (EU) No 537/2014. When the total fees received from

[75] Evidence is provided in BT-Drucksache 15/3419 (2004), BT-Drucksache 15/4054 (2004), BT-Drucksache 16/2858 (2006), BT-Drucksache 16/10067 (2008) and BT-Drucksache 16/12407 (2009).
[76] Cf. for a network definition Art. 319b HGB and Lenz (2010).
[77] Cf. BT-Drucksache 15/3419 (2004).
[78] Cf. RegE BilReG, BT-Drucksache 15/3419 (2004) and Ebke (2008).
[79] Cf. RegE BilReG, BT-Drucksache 15/3419 (2004).

a public-interest entity (PIE) in each of the last three consecutive financial years are more than 15 % of the total fees received by the statutory auditor, this fact must be disclosed to and discussed with the audit committee with respect to independence threats and possible safeguards.

Stronger rules of the Auditing Practices Board (APB) exist in the U.K. and will exist in the future even considering the new EU Regulation for PIE audits.[80] According to APB Ethical Standard (ES) 4 (Revised December 2010, no. 31) "where it is expected that the total fees for both audit and non-audit services receivable from a listed audited entity and its subsidiaries audited by the audit firm will regularly exceed 10 % of the annual fee income […] the firm shall not act as the auditor of that entity and shall either resign as auditor or not stand for reappointment, as appropriate." Insofar as profits are not shared on a firm-wide basis the reference for the 10 % benchmark is the basis on which the audit engagement partner's profit share is calculated. „Regularly" does not mean a longer time span of, for example six years like in Germany. We conclude this because APB ES 4 no. 41 (Rev. Dec. 2010; now FRC 2016 par. 4.53) stipulates the following with respect to a new audit firm for which the requirements relating to economic dependence may be difficult to comply in the short term. "Such firms would: (a) not undertake any audits of listed companies, where fees from such an audited entity would represent 10 % or more of the annual fee income of the firm; and (b) for a period not exceeding two years, require external independent quality control reviews of those audits of unlisted entities that represent more than 15 % of the annual fee income before the audit opinion is issued." APB ES 4 no. 35 (Rev. Dec. 2010) requires for audit firms for which the client importance measure exceeds 5% but will not regularly exceed 10% the consideration of appropriate safeguards to eliminate or reduce the independence threat to an acceptable level. To sum up, in the U.K. an "audit firm is deemed to be economically dependent on a listed audited entity if the total fees for audit and other audit services from that entity and its subsidiaries which are audited by the audit firm represent 10% of the total fees of the audit firm or the part of the firm by reference to which the audit engagement partner's profit share is calculated. Where such fees are between 5% and 10%, the audit engagement partner and the

[80] Cf. Turley (2008) for an overview about UK-regulations in the past, especially with respect to economic dependence.

Ethics Partner consider the significance of the threat and the need for appropriate safeguards" (APB ES 4, Rev. Dec. 2010, no. 37; similar FRC 2016 par. 4.49). The UK regulations were much stronger than the German rules and remain essentially unchanged in the Revised Ethical Standard 2016 of the Financial Reporting Council (FRC 2016).

Based on APB ES 4 (Rev. Dec. 2010) and on the Revised Ethical Standard (FRC 2016) we consider as potential critical boundaries of client importance measures 5%, 10% and 15% of the total fee income of the respective audit firm over a period of two or three consecutive years. The IESBA Code of Ethics 2012/2015 also refers to a critical limit in the specified time-period: "Where an audit client is a public interest entity and, for two consecutive years, the total fees from the client and its related entities […] represent more than 15% of the total fees received by the firm expressing the opinion on the financial statements of the client, the firm shall disclose to those charged with governance of the audit client the fact that the total of such fees represents more than 15% of the total fees received by the firm, and discuss which of the safeguards below it will apply to reduce the threat to an acceptable level, and apply the selected safeguard" (IESBA 2012, 290.222; IESBA 2016, 290.217).

With respect to fees from non-audit services, despite restrictions concerning the type of non-audit services, no explicit quantitative limits exist in Germany in the analyzed time-period. In the U.K., APB ES 5 no. 5, 6 rules: "In relation to non-audit services, the main self-interest threat concerns fees and economic dependence and these are addressed in APB Ethical Standard 4. […] In the case of listed companies where the fees for non-audit services for a financial year are expected to be greater than the annual audit fees, the audit engagement partner shall provide details of the circumstances to the Ethics Partner and discuss them with him or her." It can be concluded that the APB focuses on the relationship between non-audit fees and audit fees. According to the new Art. 4 para. 2 EU Audit Regulation, the provision of non-audit services to the audited entity, its parent undertaking or its controlled undertakings is only allowed if the total fees for such services for a period of three or more consecutive years are equal to or less than 70 % of the average of the audit fees paid in the last three consecutive financial years for the audited entity and its parent or controlled undertakings. In the Revised UK Ethical Standards 2016 fees from

network members are included (FRC 2016, 4.34R); if individual or aggregate non-audit fees are expected to be greater than the aggregate (or the individual firm's) annual fees for the audit engagement, the engagement partner must disclose and discuss this fact with the ethics partner.

Based on the presented regulations and rules, we consider 40% and 50% non-audit fees of the total fee income from the respective entity audited by the respective audit firm in two or three consecutive years as critical regulatory boundaries for the client importance measures in our study.

3.2.2 Increased Transparency with Respect to Economic Dependence from Audit Clients

In addition to the above-mentioned independence rules the German legislator has – in accordance with the Article 40 of the Directive 2006/43/EC and the Article 49 sent. 1 let. (a) no. 15 of Directive 2006/43/EC as amendment of the Article 43 sent. 1 of the Directive 78/660/EEC of the European Parliament – initiated regulations that enable outside stakeholders to evaluate the economic dependence of an audit firm.

3.2.2.1 Audit and non-audit fee disclosure

Since the financial year 2005 legal requirements (Art. 285 sent. 1 no. 17, Art. 314 para. 1 no. 8 (former) HGB) concerning the mandatory disclosure of information about audit and non-audit fees have been in place in Germany.[81] This fee information shall disclose a "too strong dependence of the auditor from audit and consulting fees"[82] of a specific client and shall enable interested parties to react in single cases if there seems to be a violation of the independence requirements in Art. 319 para. 2 or para. 3 HGB. The fee disclosure rule shall further have an influence on auditor's behavior because auditors may refrain from an audit engagement with too low fees even in a competitive audit market. Between 2005 and 2008 the disclosure applied only for capital market oriented companies for single and consolidated HGB and IFRS financial statements.[83] Since the financial year 2009 the disclosure obligation has been extended to all large limited companies with respect to

[81] Cf. Art. 58 Abs. 3 S. 1 EGHGB, BT-Drucksache 15/3419 (2004), Sultana and Willeke (2005) and IDW (2005).
[82] BT-Drucksache 15/3419 (2004, p. 25).
[83] Cf. Bischof (2006) and Lenz et al. (2006b) with respect to US-GAAP financial statements.

single financial statements (Art. 285 sent. 1 no. 17, Art. 288 HGB) and without exceptions to all parent companies, which are obliged to prepare consolidated financial statements (Art. 315 para. 1 no. 9 HGB). There is an exemption for subsidiaries if the fee disclosure is contained in the consolidated financial statements of the parent company.

According to Art. 315 para. 1 no. 9 HGB or Art. 285 sent. 1 no. 17 HGB the (parent) company has to disclose the fees which are charged for the respective financial year from the auditor of the consolidated financial statements itemized into the following four categories: (i) audit fees, (ii) fees for other assurance services, (iii) fees for tax consulting, and (iv) fees for other services. The current German rules do not require the disclosure of audit and non-audit fees, which are charged from the parent companies auditor to subsidiaries of the parent company but a statement from the German Auditor's Association (IDW 2010) recommends this disclosure in an explicit way ("thereof fees charged to subsidiaries"). If a company decides to voluntarily disclose fees charged from network members of the auditor, this should also be made explicitly. We do not discuss further interpretative problems of the disclosure rules and refer to the literature, which also discusses the diverse disclosure practice (e.g., Bischof 2006; Lenz et al. 2006b; Kirsch et al. 2013). We assume further that the recommendations of the German Auditors Association (IDW) are obeyed in practice.

3.2.2.2 Financial disclosure in transparency reports

The Professional Oversight Reform Act (Berufsaufsichtsreformgesetz, BARefG) of September 2007 implemented Art. 40 (Transparency Reports)[84] into the German law (Art. 55c WPO old version).[85] Since April 2008 transparency reports of audit firms are accessible over the homepage of the WPK or the respective audit firm itself. The aim of the German government was to give public information about "ownership, governance and quality structures of audit firms."[86] By this means, public trust into auditor's work and into the oversight of the profession should be strengthened (e.g., Naumann and Hamannt 2007). The German Chamber of Auditors (WPK) and the German Auditor Oversight

[84] Cf. EU Directive 2006/43/EG (2006).
[85] Cf. Pfitzer et al. (2007) and BT-Drucksache 16/2858 (2006).
[86] BT-Drucksache 16/2858 (2006)

Commission (AOC; APAK) can use transparency reports as a source for potential disciplinary actions.[87] According to Art. 55c para. 1 sent. 1 WPO auditors, who have at least one mandatory audit of a public interest entity (i.e. capital market oriented limited liability companies; Art. 319a HGB; Art. 264d HGB) have to disclose a transparency report within three months after the end of the fiscal year in which the audit was performed.[88] For the purpose of our study, the following information items of transparency reports are of interest:

- A *list of companies of public interest* for which a mandatory audit has been performed in the preceding financial year (Art. 55c para. 1 sent. 2 No. 5 WPO).
- *Firm level financial information* according to Art. 55c para. 1 sent. 3 No. 3 WPO in conjunction with Art. 285 no. 17 HGB, i.e., the auditor or audit firm has to disclose in the transparency report the total annual fees from all clients split into the following categories: audit fees, fees from other assurance services, fees from tax consulting, fees from other services.

3.3 Economic independence, audit and financial reporting quality

In their extensive review, DeFond and Zhang (2014) emphasize that high audit quality is not the discrete verification of no GAAP violations in the financial statements (e.g., DeAngelo 1981a), but the assurance of a high quality financial reporting. Gaynor et al. (2016) affirm that even a correct audit opinion might not entail high audit quality, if the underlying audit performed by the auditor was insufficient. In this respect, the PCAOB (2013) defines audit quality as independent, reliable audits and robust audit committee communications to meet the criteria of investors, which highlights the demand for high-quality audit processes, especially certain input factors such as engagement team competency, experience, audit firm's internal system of quality control and, independence of the auditor or audit office. Consequently, DeFond and Zhang (2014) elaborate that, based

[87] Cf. Schnepel (2013) and Art. 55c WPO, pt. 2.
[88] Cf. Schnepel (2013) and Art. 55c WPO, pt. 5.

on high-quality audit processes, pre-audited financial reporting quality, which is determined by the firm's financial reporting system as well as innate characteristics such as the fit of the GAAP for mapping the firm's assets into financial statements (e.g., Dechow et al. 2010), will be enhanced, particularly when pre-audited financial statements are of low quality. What is important, consistent definitions of financial reporting quality lack as well as holistic approaches to capture the degree of quality. When characterizing financial reporting quality, related literature (e.g., Dechow et al. 2010; Gaynor et al. 2016) uses conceptual accounting frameworks. They see financial reporting as a way of providing decision-useful information to existing and future stakeholders. In more detail, the presented information requires relevance, e.g., influencing the decisions of stakeholders by confirming the past or predicting the future, and reliability, e.g., being free from error, neutral and complete. Because of the high complexity of financial reporting quality, several dimensions exist (e.g., Dechow et al. 2010). Gaynor et al. (2016, p. 2) define higher quality financial reports as those that "*are more complete, neutral, and free from error and provide more useful predictive or confirmatory information about the company's underlying economic position and performance*."

In our study, we rely on client independence of the auditor as one major input determinant of the audit process, ultimately audit quality, and study the influence on EM as one major financial reporting quality measure.[89] EM has been extensively used and captures a firm's opportunistic incentives to influence stakeholders' decisions, thereby decreasing the quality of financial reporting due to intentional decisionism. Prior EM studies two different types, AEM and REM. Accrual-based activities change methods or estimates in financial statements (e.g., DeFond & Subramanyam 1998; Dechow et al. 2003; Kothari et al. 2005; Dechow et al. 2010) after fiscal-year end without altering cash flows. On the other hand, recent literature focuses on possible real activities of earnings manipulation (e.g., Roychowdhury 2006; Cohen et al. 2008; Cohen & Zarowin 2010; Badertscher 2011; Zang 2012), which reflect alterations of cash-flow-affecting business processes prior to fiscal-year end that ultimately change the presented earnings in the financial reporting. Besides auditor competence, auditor independence is an important component of audit quality,

[89] DeFond and Zhang (2014) categorize financial reporting quality as an output-based audit quality measure, while we follow Gaynor et al. (2016) and regard financial reporting and audit quality as different constructs.

which influences the possibilities of AEM that the management of the entity can exert. However, this restricting effect may also indirectly affect the use of the other kind of EM, e.g., REM. In particular, although real activities might not be a part of audited reporting quality insofar as they are correctly represented in the financial statement according to the applicable accounting standards, managers might exert a switch from AEM to REM when AEM is limited, likewise causing deteriorated earnings numbers.

3.3.1 Auditor size and economic dependence from clients

The relation between auditor size and his/her potential financial dependence has been formulated in a simple and idealized way (e.g., DeAngelo 1981b, Simunic 1984, Reynolds & Francis 2001, Chung & Kallapur 2003, Larcker & Richardson 2004, and Ewert 2004). The financial importance of a client increases with increasing size of the client in terms of the present value of future rents as difference between audit fees and marginal audit costs. Likewise, it increases with decreasing size of the audit firm or office in terms of the present value of total rents from all clients less total audit costs of the audit firm or office. As this discrepancy between auditor and client size enlarges, potential incentives for auditors evolve to report favorably about the client, concede in negotiations about reporting decisions and issue going concern opinions. Decisively, auditing firm's service line also plays an important rule, since high dependence on a client's audit fees might not comprise independence of the auditor when the audit service line is less important for the overall firm's sales. On the other hand, economic dependence, especially for publicly listed clients, may also impose greater audit risk in terms of possible reputation and litigation costs associated with audit failures, controversies or financial distress of the client. Reputation and litigation risk might be even more pronounced if the auditor's reputation in the market segment is undetermined, e.g., for smaller auditors with few publicly listed clients and high experience only in the audit of non-publicly listed clients. These opposing effects shape an ambiguous idea about the interaction between the two parties (e.g., Tepalagul & Lin 2015).

For empirical aims, additional premises are necessary. We cannot observe rents (profits) resulting from audit and non-audit services. Therefore, we assume a linear relation between rents and auditor's revenues. Furthermore, most legal professional requirements

also define fee dependence as percentage of the total revenues of an audit firm. Especially, the analytical considerations for audit firm size's influence on auditor independence should be empirically tested because it follows not conclusively that small and midsized audit practices are more economically dependent from their clients than larger audit firms are. Depending on the client and service structure of an audit firm, it is empirically possible and probable that smaller audit firms with respect to economic dependence do not exceed certain critical thresholds of 5 %, 10 % or 15 %. Therefore, an empirical analysis about the relation between audit firm size und economic dependence is necessary in this audit market segment. Most previous studies do not measure economic dependence of auditors in accordance with the above considerations.[90] Notable exceptions are Chung and Kallapur (2003) for the U.S. audit market and Quick and Sattler (2011a,b), Lopatta et al. (2015) for Germany. Due to lacking data availability, even more recent studies, e.g., Hope and Langli (2010), Sharma et al. (2011), Coulton and Ruddock (2012) or Svanström (2013) only use fee data from publicly listed clients and do not use the total fees of an audit firm from public and private clients in the denominator of their measure of client importance. In our study that would lead to a severe overstatement of the economic bonding of smaller auditors because these auditors regularly have only few listed clients.

3.3.2 Theoretical framework

Figure 4 shows our theoretical framework. Box I shows in a stylized way the responsible parties for the appointment of an auditor. They determine quantity, type and fees of the demanded audit and non-audit services. In the German context, the auditor is appointed by the shareholders at the annual meeting based on the supervisory board's proposal (Art. 119 para. 1 no. 4; Art. 124 para. 3 sent. 1 AktG). For PIEs with an audit committee[91], the proposal rests on a recommendation by the audit committee (Art. 124, para. 3, sent. 2 AktG). The supervisory board and in particular the audit committee, as the representative body of the shareholders, are responsible for the negotiation of the audit fees with the

[90] E.g., Frankel et al. (2002) use the relation between non-audit fees to total fees of an audit client, the rank transformation of audit and non-audit fees and absolute value of total fees to measure economic dependence. Craswell et al. (2002) use the relation of audit fees to total fees of a client. Zimmermann (2006) measures client dependence with the absolute (logarithmic) values of non-audit and audit fees and the relation non-audit fees to total fees.
[91] The establishment of an audit committee is voluntary. If no audit committee exists, the tasks of the audit committee are taken over by the whole supervisory board (Art. 107 AktG).

auditor. The executive part of the board, e.g., CEO and CFO, is responsible for the preparation of the financial statements and may decide to order non-audit services from the auditor when permission is granted by the supervisory board or the audit committee.

Figure 4 – Theoretical framework for economic bonding and auditor independence

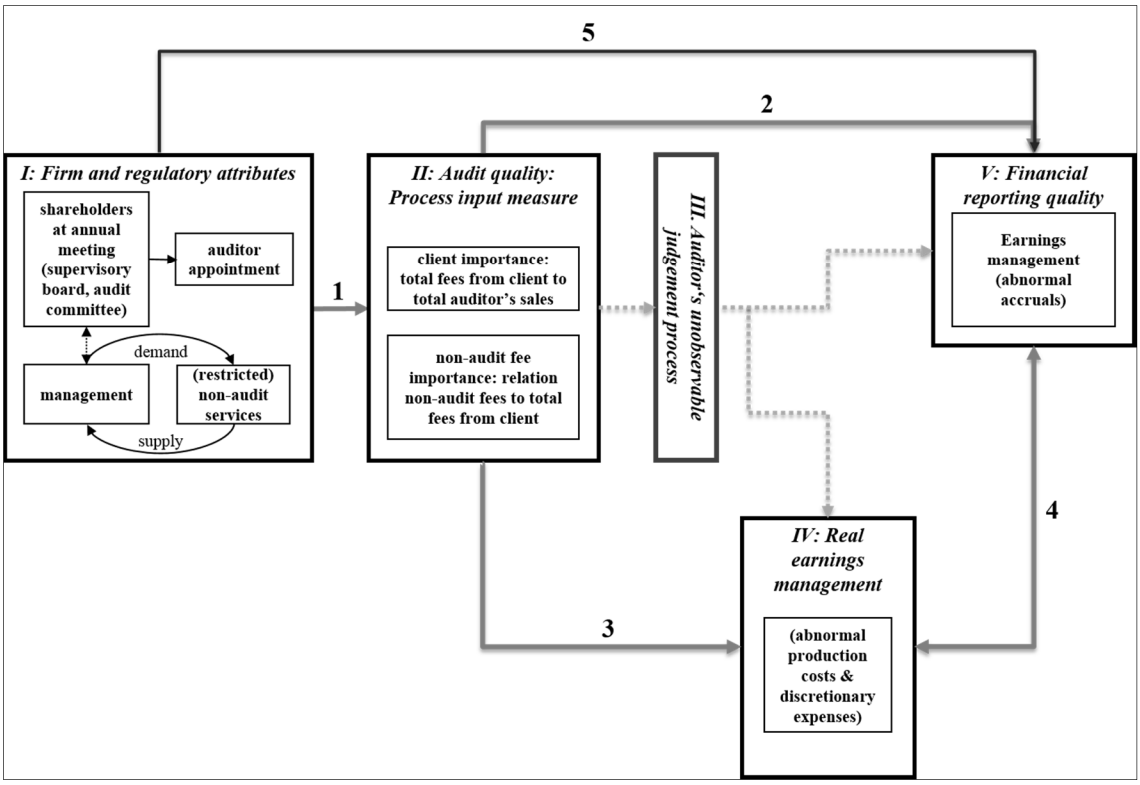


Figure shows the theoretical framework for the underlying analysis.

The engagement process of the auditor creates an economic bond between auditor and client (arrow 1 to Box II). The appointed auditor, being an (economic) agent of the firm, will consider the principal's interest based on the resulting economic dependence. On this occasion, client's and auditor's incentives, e.g., for the client the increase in reliability and decrease in moral hazard concerns with stakeholders and for the auditor litigation and reputation risk, play an important role. On the other hand, more opportunistic incentives might exist on both sides, e.g., for the manager to use income-increasing or -decreasing EM and for the auditor to retain the client's fees (e.g., DeFond & Zhang 2014). Our measures for economic dependence, e.g., client importance and non-audit fee importance, are two major audit process input measures.

Dependent on the set of audit process input measures, Box III symbolizes the unobservable information process of the auditor or the audit team. A strong economic bond may impair auditor's independence in the professional judgment process (dotted arrow between Box II and Box III), resulting in a lower audit quality. In particular, when the auditor's opinion is not congruent with his/her inner opinion about the true state of the financial statements, the management might try to make use of the economic bond and influence financial reporting. Decisively, auditor's professional judgment has a significant influence on the quality of the published financial statements, i.e., the extent of AEM (arrow 2).

However, if the interests of the supervisory board are perfectly aligned with shareholders' interests, then there should be no impaired audit judgment due to economic bonding because the supervisory board effectively controls auditor's appointment and remuneration, restricts non-audit services and implements an effective audit committee. Therefore, even in the case of a strong economic bond – created by a well-knowing supervisory board – the auditor may have incentives to deliver a high audit quality in order to become appointed again by the supervisory board while being controlled effectively by the audit committee.

Arrow 5 symbolizes the direct influence of firm and regulatory attributes on financial reporting quality, e.g., the existence of a strong and independent supervisory board and/or audit committee. Consequently, if an independent auditor effectively restricts AEM, client's management may evade to REM, since auditors' engagement does not explicitly involve the limitation of this kind of EM. Therefore, we control for a possible trade-off between AEM and REM (arrow 4). There are also direct effects between audit and non-audit fees and REM (arrow 3). For example, the auditor may increase audit fees for clients with aggressive EM because they evaluate that as risky behavior and therefore increase their audit effort and/or the risk premium in audit fees (Greiner et al. 2017). Clients who use REM to meet earnings targets may also have an increased demand for consulting services.

3.3.3 Related literature and hypotheses development

Tepalagul & Lin (2015) present an extensive literature review about auditor independence and audit quality. They highlight the focus on the relationship between client importance, mostly proxied by non-audit services, and financial reporting quality, proxied by AEM (Frankel et al. 2002; Ashbaugh et al. 2003; Ferguson et al. 2004; Chung & Kallapur 2003; Sharma et al. 2011; Chi et al. 2012), modified audit opinions (Reynolds & Francis 2001; DeFond et al. 2002; Geiger & Rama 2003; Hope & Langli 2010) or earnings restatements (Raghunandan et al. 2003; Kinney et al. 2004).

An overview of relevant prior studies for the U.S. (10 studies), U.K. (6 studies), Australia (6 studies) and New Zealand (4 studies) can also be found in Sharma et al. (2011). For a meta-analysis about the relationship between audit quality, corporate governance and EM Lin and Hwang (2010) identify 49 prior studies. Sattler (2011, Table 9-2, pp. 240-289) delivers a detailed discussion of 19 empirical studies.

Interestingly, there are only few studies focusing on the influence of client importance resulting from the sum of audit and non-audit fees in relation to total sales (or total sales resulting from the listed client market) of the respective audit firm on AEM. These are (in chronological order): Chung and Kallapur (2003, USA, no association); Larcker and Richardson (2004, USA, no association); Mitra (2007, USA, no association); Cahan et al. (2008, NZ, pos. association); Sharma et al. (2011, NZ, pos. association moderated by audit committee characteristics); Coulton and Ruddock (2012, Australia, no association). There are only two German studies (see for details below): Quick and Sattler (2011b, no association) and Lopatta et al. (2015, pos. association with income increasing discretionary accruals).

There are some reasons why we might not expect that a strong economic bond between auditor and client may lead to a negative effect on financial reporting quality. First, auditors are aware of the threats from impaired independence and therefore put in place safeguards like quality control and engagement reviews (Coulton & Ruddock 2012, p. 5 et seq.). DeFond and Zhang (2014) argue that especially for large clients' auditor's reputation and litigation concerns may offset threats from financial dependence but admit that

the evidence is mixed and different for discretionary accruals and GCOs. As argued above, a strong and effective supervisory board and/or audit committee should also be able to restrict and control threats from impaired audit quality due to economic bonding. In our setting – the small and midsized audit market segment – the average auditor has only a small number of listed clients. Gaining a listed firm as a client increases the reputation of the non-Big4 auditors, i.e., they should have high incentives to retain this reputation. In addition, the auditors know that they are under the spotlight of the German Auditor Oversight Commission and their clients are under the spotlight of the Financial Reporting Enforcement Panel. This should especially be the case for non-Big4 auditors, since they tend to have no matured reputation in the market for publicly listed clients. Combined, these effects may be effective in securing the audit quality despite a significant economic auditor-client bond.

Regulators conjecture (in contrast to audit literature) that an auditor is economically dependent on a specific client if the total fees for audit and other audit services from the audited entity and its subsidiaries exceed certain thresholds, e.g., 5%, 10% or 15%, in a specified time period. This may compromise audit independence and hence lead to impaired financial reporting quality. The competing theoretical arguments and the mixed empirical evidence lead us to the following (regulatory-driven) hypothesis stated in null form:

H1: Client importance for audit clients *i* of auditor *k* exceeding a certain threshold, e.g. 5%, 10% or 15% in a defined period of at least two or three consecutive years is not associated with lower financial reporting quality, i.e., higher AEM.

It should be mentioned that the German context, in contrast to the U.K., entails no legal binding of the thresholds in **H1**, i.e., the auditor may exceed the limits without the fear of negative sanctions from auditor oversight bodies. Therefore, we can use this setting to test whether in a more liberal audit environment economic bonding has in fact had negative consequences on audit quality. With **H1** we do not assume linear relationship between client importance and audit quality like most other studies (e.g., Chung & Kallapur 2003; Sharma et al. 2011; Coulton & Ruddock 2012). With the exception of Chung and

Kallapur (2003) and Larcker and Richardson (2004)[92] all other studies do not measure client importance as suggested in our study. Sharma et al. (2011) and Coulton and Ruddock (2012) admit that an inherent limitation in this line of research as a whole is that because "only publicly listed clients are required to disclose audit and NAS fees" leads to an understatement of the total fees earned by an audit firm and therefore overstates "the relative importance of a client to a partner or audit firm."

The bulk of prior studies concentrate on the effect of non-audit fees and it can be summarized that "there is no unequivocal evidence that client importance measured as non-audit fees paid to the auditor affects the quality of financial reporting" (Sharma et al. 2011). Sharma (2014) reviews 45 studies about the relationship between NAS and auditor independence and summarizes that the overall evidence is "far from conclusive". DeFond and Zhang (2014) summarize that "studies using output-based proxies find that NAS does not impair auditor quality, and some NAS may even improve it." Main controversial themes are measurement, i.e., with respect to client importance[93] and financial reporting quality, and causality issues. On the other hand, a meta-analysis of 10 studies done by Lin and Hwang (2010, Table 3, p. 66, p. 70) states that "the evidence strongly supports the negative effect on earnings quality […] of a strong economic bond between the client and the external auditors, as measured by the high fees paid for non-audit services, relative to total fees" (i.e., the non-audit fee-ratio).

A common assumption in this literature is that high non-audit fees may impair auditor independence and therefore financial reporting quality. Again, there are competing arguments that non-audit services may improve audit quality by enabling knowledge spillovers improving auditor competency and efficiency (Simunic 1984). Furthermore, potential negative effects on auditor independence may be controlled for by internal and external safeguards, e.g., internal separation of audit and non-audit work, quality controls, audit committees approval and oversight of non-audit work, reputation and litigation concerns.

[92] "As in Chung and Kallapur (2003) we obtain the total fee revenue for each auditor from *Accounting Today*" (Larcker & Richardson 2004).
[93] "However, there is no agreement on how to measure this economic bond" (Lin & Hwang 2010).

Most studies use the absolute value on non-audit fees, rank transformations, and the relation of non-audit fees to audit fees or total fees from the respective client, but seldom the relation of non-audit fees to total fees of the audit firm in conjunction with a threshold approach as we do.[94] Regulators argue that the relation between non-audit fees and average audit fees should not exceed a certain limit, e.g., 70 % according to the new EU Audit Regulation. That translates into a 41 % non-audit fee threshold[95] for non-audit fees with respect to total fees. Again, the competing theoretical arguments and the mixed empirical evidence leads us to the following (regulatory-driven) second hypothesis stated in null form:

H2: Client importance for audit clients *i* resulting from non-audit services of auditor *k* exceeding a certain threshold, e.g., 40 % or 50 % of total fees from client *i* in a defined period of at least two or three consecutive years is not associated with lower financial reporting quality, i.e., higher AEM.

Again, in the period under study, in Germany no such quantitative restrictions for non-audit services were in place. Additionally, we test other common measures, e.g., the relation of non-audit fees to audit-fees.

Since financial reporting quality as a theoretical construct is not directly observable, proxies have to be used. A common measure therefore is AEM as a direct financial statement attribute. In contrast, REM "alters normal firm operations, impacts current and future cash flows, imposes additional costs, and sacrifices firm value" (Greiner et al. 2017, p. 86, with reference to Cohen & Zarowin 2010, Badertscher 2011 and Evans et al. 2015). REM is generally not a direct financial statement attribute and therefore less scrutinized by the auditor (e.g., Chi et al. 2011; Kim & Park 2014). To the best of our knowledge, there is only one study, which examines the association between audit quality and REM (Chi et al. 2011), but no study analyzing the relationship between auditor's economic dependence on the client and REM. Chi et al. (2011) show for a sample of firms with strong incentives

[94] We additionally use the relation of non-audit fees to audit fees as well as the absolute value of non-audit fees in our robustness tests.

[95] $TotalFees\ (TF) = AuditFees(AF) + NonAuditFees\ (NAF)$ and $NAF = 0.7 \cdot AF$; $NAF = 0.7 \cdot (TF - NAF) \rightarrow \frac{NAF}{TF} = \frac{0.7}{1.7} \approx 0.41$

to manage earnings upwards that higher quality auditors (proxied by auditor industry expertise and audit firm size, i.e., BigN versus non-BigN) restrict AEM and, therefore, clients resort to REM.

If there is a tradeoff between AEM and REM (Zang 2012) and generally, REM is costlier than AEM, then clients of more dependent auditors should prefer AEM and clients of independent auditors, who restrict AEM would switch to REM if there were enough incentives for them. However, since the relation between economic bonding and AEM is inconclusive, i.e., if there is no association between economic bonding and AEM, we expect also no association between economic bonding and REM. Therefore, we formulate the hypothesis **H3** in null form:

H3: Ceteris paribus, there is no association between auditor-client economic bond and REM.

In general, auditors should be able to identify REM because a business risk-oriented audit approach requires a thorough understanding of the business strategy and the related operational management decisions (e.g., Greiner et al. 2017). Commerford et al. (2016) show in a qualitative interview-based study that auditors are aware of REM and express concerns about REM. They respond by increasing engagement risk assessments and altering audit procedures. Kim and Park (2014) find that REM is associated with auditor resignations and Greiner et al. (2017) show that aggressive income-increasing REM is positively associated with both current and future audit fees, whereas the economic impact on (current and future) audit fees is low (1.57 % versus 1.53 %).

Prior German evidence

Quick and Sattler (2011b) analyze whether client importance, measured as the fraction of total fees from a client to the total revenues of the audit firm, has an influence on discretionary accruals comprehended from the performance-adjusted Jones model. They use a pooled sample comprising 329 consolidated IFRS financial statements of firms (excluding banks and insurance companies) listed in the German Prime Standard at the respective year-end of 2005-2007. No significant statistical relation can be identified but it should be noted that this specific audit market segment is heavily dominated by Big4 auditors

(market share measured by number of clients between 77% and 83% within the three years) and generally Big4 auditors do not heavily depend on single clients. In our study, we focus on Non-Big4 auditors.

Using essentially the same sample, Quick and Sattler (2011a) follow Sattler (2011) and analyze for the German audit market if non-audit services impair auditor independence. They apply different fee ratio variables, e.g., the ratio of non-audit fees or tax consulting fees to total fees, to measure an impaired independence resulting from non-audit services. An OLS regression with 341 firm-year observations shows that an increasing tax consulting fee ratio increases positive (income-increasing) discretionary accruals, while increasing other non-audit fees[96] increase AEM in terms of the absolute and positive discretionary accruals. The authors conclude that particularly other non-audit fees may impair auditor independence. They additionally test certain thresholds for non-audit fees or categories of non-audit fees (> 10 % or > 25 % of total fees) and their effects on AEM. The results show a statistically positive impact on the absolute discretionary accruals, while this effect seems to be driven by other non-audit fees. In addition, exceeding the 10 % limit of other non-audit fees leads to higher positive discretionary accruals (e.g., Quick & Sattler 2011a).

The studies together as well as international evidence seems to imply that it may be appropriate to differentiate between effects on discretionary accruals of different categories of non-audit fees despite the fact that we currently have no clear theory about the underlying reasons. With respect to non-audit fee client importance, there should be no difference between the categories but it may be argued that, for example, other assurance services, e.g., reviews of interim reports, which are closely related to audit fees can lead to knowledge spillovers and therefore increase audit efficiency, but will particularly not distract the auditor from audit objectives. We also have no precise theoretical understanding of why the effects for absolute, positive (income increasing) or negative (income decreasing) discretionary accruals should be different. While absolute values assume both kinds of EM as comparably erroneous and detectable for the auditor, possible reasons for inconsistent results on signed discretionary accruals could include unequal distribution of

[96] Other non-audit fees means fees other than assurance and tax consulting fees, i.e. fees for M&A consulting.

positive and negative values. E.g., excessive insistence on strict financial reporting by the auditor could restrict income-increasing EM to such a high degree that all discretionaries from accruals models turn more negative. Altogether, interpretation of the Quick and Sattler (2011a) results remains inconsistent.

The sample of Lopatta et al. (2015) is comprised of 840 firm-years from listed firms (without banks and insurance companies and firms domiciled in foreign countries) from the premium segments of the German stock market over the period 2005-2011. They conjecture a positive relationship between economic dependence – measured as total audit client's fees to total auditor's revenues – and the level of AEM proxied by discretionary working capital accruals estimated using the performance-adjusted modified Jones model. They further assume a positive relation between the percentage of non-audit fees to total client's fees and AEM. Their regression results cannot reject the hypotheses with respect to positive discretionary working capital accruals, i.e., if the audit client uses income-increasing AEM. Total audit firm's revenues in Lopatta et al. (2015) stem from Lünendonk market studies based on voluntary audit firm's disclosure. In our opinion this source is not really trustworthy because (i) Lünendonk reports mostly voluntary disclosed fees for national networks of audit firms and (ii) in several cases this is not the appropriate basis for matching fees from a specific audit client with total network turnover, e.g., if the network is not the relevant economic entity for measuring economic bonding.

3.4 Research design

3.4.1 Sample

Table 24 summarizes the sample selection procedure. We use an initial dataset from 2007-2014 of 2,441 firm-years containing all German publicly listed clients (i.e., capital market oriented limited liability companies; Art. 319a HGB; Art. 264d HGB) of non-Big4 auditors.[97] We hand-collect financial information of the non-Big4 auditors from the respective

[97] We keep each client in the sample as long as it has been audited by a non-Big4 auditor at least once within the sample period.

transparency reports 2009-2016[98]. We use Thomson Reuters Datastream to collect financial information of the audited entities.[99] Furthermore, we use the financial reports of the audited entities to collect sufficient data concerning economic bonding, audit engagement and corporate governance controls. We exclude all financial institutions due to disparate balance sheets and ways to pursue EM, which accounts for a loss of 104 observations. Consequently, we account for a loss of 1.097 firm-years due to insufficient financial information for the estimation of our EM models. We further account for a loss of 111 firm-years due to missing data on economic bonding and 47 firm-years due to missing data on control variables. In addition, we require each industry-year combination to account for at least 10 observations, resulting in a loss of 30 firm-years. Subsequently we use an unbalanced panel with adequate information for all dependent and independent variables consisting of 1.052 firm-year observations.

Table 24 – Sample selection procedure

Selection mode	Number of observations
All PIEs with a non-Big4 auditor with available financial information in transparency reports from 2009-2016 (2007-2014)	2,441
Less:	
Financial institutions	104
Missing data in Thomson Reuters Datastream for estimation of EM models	1,097
Missing data in financial reports on economic bonding	111
Missing data in financial reports on control variables for corporate governance & audit engagement	47
Firm-years with less than 10 industry-year firm years available	30
Final Sample: Adequate information for all independent (economic bonding & controls) and dependent (EM) variables	1,052

3.4.2 Measurement of important variables and empirical design

3.4.2.1 Measurement of financial reporting quality

Although DeFond and Zhang (2014) regard EM as an output-based accrual quality measure, we distinguish audit and financial reporting quality following Gaynor et al. (2016) and see EM as a direct measure of financial reporting quality rather than an indirect measure of audit quality. As already discussed in chapter 2.2, we use the cross-sectional and

[98] Financial information in the transparency reports is presented for the preceding financial year, while this financial data includes fees from the preceding financial year of the client. Therefore, using transparency reports from 2009-2016, we are able to analyze a publicly listed client dataset from 2007-2014.
[99] In particular, we use Thomson Reuters Datastream to collect financial information of the PIEs to enable inclusion of control variables for market-based size and market-to-book ratio.

pooled modified, performance-matched Jones model (e.g., Subramanyam 1996; DeFond and Subramanyam 1998; Dechow et al. 2003; Kothari et al. 2005) to comprehend AEM, while we apply the cross-sectional and pooled models for REM proposed by Roychowdhury (2006), standardize each of the proxies and calculate a combined signed measure REM_{it} (e.g., Cohen et al. 2008; Greiner et al. 2017).[100]

To study apply they the trade-off relation between AEM and REM in more detail, we introduce a separate decile-based proxy following Bozzolan et al. (2015). We therefore compute deciles for AEM ($DACC\_abs_{it}$) and REM (REM_{it}) to comprehend the trade-off proxy $REMvsAEM_{it}$ as the ratio of the REM_{it} decile to the sum of the two deciles for AEM and REM.[101]

3.4.2.2 Measurement of audit quality & measurement issues

We have to transform theoretical concepts into valid and reliable operational measures. In general, measurement issues and the selected statistical method have a material effect on the empirical result and therefore should be considered carefully.

When studying the heterogeneity of our sample, we find several probable measurement issues concerning economic bonding. First, the determination of economic units with Non-Big4 audit entities can be challenging due to interrelations between the entities. This can have a considerable influence on the resulting economic bond used in the respective study. Likewise, it might affect the decisions of regulators with respect to intolerable auditor dependence. E.g., the audit firm Dr. Kleeberg & Partner GmbH holds 100% of the shares of the Crowe Kleeberg Audit GmbH. Both the parent and subsidiary have audit mandates in our sample. When measuring economic bonding, it is debatable whether the total audit sales of the subsidiary are relevant for the mandates of the subsidiary, or whether economic bonding should rather be measured based on the total sales of the whole audit group.[102] Likewise, the audit firm Crowe Horwarth Deutschland GmbH has only one PIE client and the transparency report contains only the fees from this client,

[100] We follow the calculation proposed by Kim et al. (2012), which is also used in chapter 4: $REM_{it} = CFO\_abn_{it} - PROD\_abn_{it} + DISEXP\_abn_{it}$. This proxy for real activities EM is negatively scaled and standardized.
[101] To demonstrate the validity of our results, we use alternative proxies for AEM and REM. We discuss the results in section 3.7.
[102] Another related problem occurs when there is a change in audit mandates between the parent audit firm and the subsidiary. Measuring economic bonding as the ratio of audit fees to total sales of the entity performing the audit, there might be a major increase in the dependence proxy when in reality there is only a shift between parent and subsidiary audit firm.

e.g., without further considerations we would conclude that economic dependence is 100 percent. A further look reveals that this company is a joint venture of two other audit firms (HSA Horwart GmbH and RWT Horwarth GmbH), has no own personnel and therefore uses the personnel of the audit firms from the RWT group and HSA Horwarth GmbH. Naturally, using the considerably higher group revenues to calculate economic dependence results in a significant change of the economic dependence proxy.

What is more, we find constellations where several audit firms bundle their activities with respect to PIE clients in a joint legal entity without own personnel (see for example PKF Deutschland GmbH). This legal audit entity is the auditor of the PIE but in conducting the audit this entity uses personnel from one of its shareholder audit firms. Whenever we are able to identify the audit firm, which in effect has performed the audit, we use the revenues of this entity to measure economic dependence.

Subsequently, we always try to identify the controlling audit entity on the supply side to measure economic dependence from audit clients as precise as possible. Therefore, we apply the audit firm group revenues as relevant denominator whenever the parent audit firm controls the subsidiary.[103]

When dealing with information published by the audit firms, another problem arises. When date of fiscal year end of the audited company is December 31 of the year, financial information of the auditing company two years forward is relevant, since for example the transparency report published in March 2016 includes sales of the year 2015, i.e., revenues from audits for the financial year 2014. Whenever fiscal year end of the audited company lies considerably before that date, i.e., end of June 2015, usage of the financial information of the transparency report two years ahead might bias the proxies, since the actual audit fee has already been paid before, which means the transparency report one year ahead is relevant. We paid attention to this bias during the collection of our data in the best possible way.

[103] However, the economic dependence could still be biased in cases where there is no sufficient information about parent-subsidiary constellations. While these cases are very rare, they could work in favor or against our hypotheses.

3.4.2.3 Empirical model

To perform multivariate tests on our hypotheses, we apply the following empirical model to test the influence of client importance from the importance of total sales of the client on the separate types of EM, while controlling for the confounding nature of several variables:

Equation 24

$$
\begin{aligned}
DACC\_abs_{it}/TA_{it-1}\ (REM_{it}/TA_{it-1}) &= \alpha_0 + \boldsymbol{\beta_1 CI_{it}} + \boldsymbol{\beta_2 DEP_{it}} + \boldsymbol{\beta_3 CI_{it} * DEP_{it}} \\
&+ \beta_4\, REM_{it}/TA_{it-1}\,(DACC\_abs_{it}/TA_{it-1}) + \beta_5 Size_{it-1} + \beta_6 MB_{it-1} \\
&+ \beta_7 LOSS_{it} + \beta_8 Litigation_{it} + \beta_9\, ACC_{it-1}/TA_{it-1} \\
&+ \beta_{10}\, CFO_{it}/TA_{it-1} + \beta_{11} Leverage_{it-1} + \beta_{12}\, \Delta TA_{it}/TA_{it-1} \\
&+ \beta_{13}\, \Delta EARN_{it}/TA_{it-1} + \beta_{14} ROA_{it} + \beta_{15} FirmAge_{it} \\
&+ \beta_{16} PartChange_{it} + \beta_{17} AudChange_{it} + \beta_{18} AudOpClean_{it} \\
&+ \beta_{19} BoardSize_{it} + \beta_{20} BoardInd_{it} + \beta_{21} VioDCGK_{it} + Year \\
&+ Industry + \varepsilon_{it}
\end{aligned}
$$

where $DACC\_abs_{it}$ is the absolute discretionary accruals and REM_{it} is the combined signed standardized measure for REM of firm i in year t.[104] CI_{it} is either client importance from sales (CIS_{it}), calculated as ratio of total fees received from client i to the total sales of the audit firm from all services, or client importance from client's non-audit fee importance, calculated as ratio of non-audit fees to total fees received from the client ($CINAS_{it}$). DEP_{it} is an indicator variable, which is 1 for firms where certain CI thresholds are exceeded, e.g., 5 %, 10 % or 15 % client importance from total fees and 40 % or 50 % client importance from non-audit fees are exceeded for two or three consecutive years.[105] $CI * DEP_{it}$ is an interaction effect variable capturing distinct effects of client importance for the high dependence part of the sample. We include either $DACC\_abs_{it}$ or REM_{it} as separate control in **Equation 24** to capture the substitution effects between the two types (e.g., Cohen et al. 2008; Zang 2012).

[104] See chapter 2.2 for a detailed discussion of the measurement of the applied EM proxies.
[105] To complement the time perspective of the independence threat, we additionally analyse whether exceeding the high dependence threshold in a single year already has a significant outcome.

We incorporate various confounding control variables to minimise omitted variable biases. While inclusion of collider or mediator controls might also cause type I errors of false rejection of the null hypothesis, we only use well-established controls from prior literature and assume them to be confounding (e.g., Gow et al. 2016).[106] We therefore follow Roychowdhury (2006) and include market-based size and growth control. EM literature has found that these influencing factors can help explain considerable variation in the discretionary proxies (e.g., Frankel et al. 2002; Gunny 2010; Sharma et al. 2011; Abbott et al. 2016). $Size_{it-1}$ is the natural logarithm of lagged market value of equity, whereas MB_{it-1} is market to book value of equity. $LOSS_{it}$ is an indicator that is one if earnings are negative, zero otherwise, hereby considering the adverse EM incentives of these situations (e.g., Hope & Langli 2010; Sharma et al. 2011; Abbott et al. 2016). $Litigation_{it}$ is an indicator control for litigation risk (e.g., DeFond & Subramanyam 1998; Cohen et al. 2008; Sharma et al. 2011). ACC_{it-1} and CFO_{it} are included to decrease influences of EM measurement errors caused by the first-stage regressions (e.g., Kang & Sivaramakrishnan 1995; Young 1999; Ashbaugh et al. 2003; Ferguson et al. 2004; Prawitt et al. 2009). In addition, we follow Kim and Park (2005) and include $Leverage_{it-1}$ as another regressor, since the degree of leverage can likewise shape EM opportunities (e.g., Ferguson et al. 2004; Li 2009; Abbott et al. 2016).

Following Chi et al. (2011) and Ali and Zhang (2015), we consider two measures for business growth. Total asset growth (ΔTA_{it}) captures the changes in assets, which influence the magnitude of total accruals due to correlation with firm's working capital. $\Delta EARN_{it}$ on the other hand controls for the change in earnings from year t-1 to t. ROA_{it} is earnings before extraordinary items divided by lagged total assets, allowing and controlling for possible non-linearity of the performance influence in the second-stage, complementing the already performance-matched discretionary accrual approach in the first-stage.[107] We capture differences in the accounting behaviour for firms with different maturation by implementing $FirmAge_{it}$ as the natural logarithm of the age of the firm, where age is calculated as 1+ year of incorporation to the Thomson Reuters database minus the

[106] However, we address the problem by testing for the robustness of our results when using a baseline model, which only includes control variables where we can cancel out mediation or collision.
[107] See chapter 2.1.4.4 for a discussion of possible non-linearities of performance for the banking models. We regard this issue as comparably important in the industry models.

fiscal year (e.g., Prawitt et al. 2009; Kim et al. 2012; Abbott et al. 2016). We additionally try to disentangle certain auditing process characteristics, which could have a significant influence on the magnitude of EM. Therefore, $PartChange_{it}$ is an indicator for the change in one or both of the partners that manage the audit client. Additionally, $AudChange_{it}$ is an indicator for a change in the audit firm. Audit quality research has shown that both changes in partners and audit firms can have significant influences on the quality of the earnings numbers (e.g., Ferguson et al. 2004; Antle et al. 2006; Lopatta et al. 2015). Likewise, prior literature includes auditor's opinion, which captures influences from opportunistic EM whenever the opinion is not clean (Larcker & Richardson 2004; Antle et al. 2006; Lopatta et al. 2015). We therefore add $AudOpClean_{it}$, indicating whether there has been a clean opinion for the financial year.

We also control for Corporate Governance impacts on the decisions about EM. We therefore include $BoardSize_{it}$, which captures the numbers of members on the board and is a common control for Corporate Governance (e.g., Bradbury et al. 2006; Marra et al. 2011; Badolato et al. 2014). Accounting research has extensively studied the question of whether larger boards can increase monitoring through more experience, knowledge and better environmental bonds. The majority of the results show that firms with a larger board can increase monitoring and therefore restrain EM (e.g., Dalton et al. 1999, Klein 2002). Alternatively, Jensen (1993) claims that larger boards are less effective since they have a lot of coordination effort and processing problems (e.g., Dechow et al. 1996; Peasnell et al. 2005). We further include $BoardInd_{it}$, measured as the supervisory board's magnitude compared to the total board size. Various studies report significant outcomes indicating that higher amounts of independent members on the board can help to increase monitoring, hence decreasing the degree of EM used (e.g., Fama & Jensen 1983; Beasley 1996; Klein 2002; Bradbury et al. 2006). Another important measure for Corporate Governance is $VioDCGK_{it}$. Art. 161 of the German Stock Corporation Act (Aktiengesetz) obliges every capital market oriented German corporation to declare either compliance with recommendations of the German Corporate Governance Codex (Deutscher Corporate Governance Kodex (DCGK)) or may choose to deviate from them, but then they are obliged to disclose deviations on an annual basis ("comply or explain-principle"). The DCGK contains recommendations on how corporations can implement

and maintain good Corporate Governance. These recommendations include establishment of an audit committee as well as collocation of the board, requirements for the compensation of board members etc. Therefore, good Corporate Governance in the shade of few or none violations on the DCGK entails a favourable auditing process, aligned with Larcker and Richardson (2004), thereby limiting the use of EM.[108]

We additionally use the following specification to study the trade-off between AEM and REM, while we exclude the control for the respective other type of EM (β_4 in **Equation 24**):

Equation 25

$$\begin{aligned} REMvsAEM_{it} = {} & \alpha_0 + \boldsymbol{\beta_1 CI_{it}} + \boldsymbol{\beta_2 DEP_{it}} + \boldsymbol{\beta_3 CI_{it} * DEP_{it}} + \beta_4 Size_{it-1} \\ & + \beta_5 MB_{it-1} + \beta_6 Loss_{it} + \beta_7 Litigation_{it} + \beta_8 ACC_{it-1}/TA_{it-1} \\ & + \beta_9 CFO_{it}/TA_{it-1} + \beta_{10} Leverage_{it-1} + \beta_{11} \Delta TA_{it}/TA_{it-1} \\ & + \beta_{12} \Delta EARN_{it}/TA_{it-1} + \beta_{13} ROA_{it} + \beta_{14} FirmAge_{it} \\ & + \beta_{15} PartChange_{it} + \beta_{16} AudChange_{it} + \beta_{17} AudOpClean_{it} \\ & + \beta_{18} BoardSize_{it} + \beta_{19} BoardInd_{it} + \beta_{20} VioDCGK_{it} + Year \\ & + Industry + \varepsilon_{it} \end{aligned}$$

$REMvsAEM_{it}$ is a decile-based variable that captures the trade-off between AEM and REM (e.g., Bozzolan et al. 2015). We split both $DACC\_abs_{it}$ and REM_{it} distrubutions into deciles and calculate $REMvsAEM_{it}$ as the decile for REM_{it} in relation to the sum of the deciles for both $DACC\_abs_{it}$ and REM_{it}. We estimate **Equation 24-Equation 25** using OLS with two-way clustering at the firm and time level (e.g., Petersen 2008; Gow et al. 2010).[109]

[108] We additionally address the importance of a best practice audit committee in chapter 3.7.4 and study the moderating effect on our initial results.

[109] Since the standard errors of our regressions may be influenced by cross-sectional and serial correlation, we apply two-way clustering by firm and year. Petersen (2008) and Gow et al. (2010) show that when both sources of correlation are existent and considerably high, sole parametrical formulation, e.g., inclusion of firm or year dummies, might not solve the problem whenever the firm and/or time effect is not fixed. Two-way clustering accounts for non-fixed effects and remains unbiased as long as the number of clusters is sufficiently high. We assume our estimation procedure with 236 firm and 6 year clusters to be appropriate. However, to address the problem of too few clusters, we apply the Cameron et al. (2011) adjustment in robustness analyses. Using industry instead of firm clustering yields no estimation advantage compared to simple one-way clustering, since only 8 industry clusters exist and these effects might be nested in firm clusters. We additionally include time and industry fixed effects to control for an industry and time fixed effect. Additional firm fixed effects would significantly decrease estimation efficiency due to high number of firms compared to the total number of observations. On top of that, we assume our hypothesized effects to be between firms.

3.5 Descriptive statistics and univariate analysis

Table 25 reports descriptive statistics for all relevant variables. We winsorize at the top and bottom 1 percent of the respective distribution to account for outliers. The mean values of signed discretionary accruals and real activities manipulation account for -0.1 and 0.3 percent of lagged total accruals. They are not significantly different from zero, indicating a favorable setting with no systematic upward or downward bias of the EM proxies.[110] The sample firms are relatively young with a mean firm age of 2.449 years and have comparable market-based size with a mean natural logarithm of market value of equity of 10.829 and an interquartile range of only 1.99. However, expected growth opportunities differ significantly with a mean market-to-book ratio of 1.746 and a standard deviation of 2.219. Compared to German settings with mainly large audit firms (e.g., 82.4 percent Big4 companies in Lopatta et al. 2015), our sample firms account for a loss in 29 percent of the years (14.3 percent in Lopatta et al. 2015), while around every third firm is located in a high-litigation risk industry. What is more, we have considerable discontinuity concerning the engagement and review partners (change of at least one partner in 41.2 percent observations), while audit firms change in 13.4 percent of the years. Corporate governance variables show a board size of on average seven members, while 65.2 percent of them are independent. Sample firms account for 6.516 violations of the DCGK, while Lopatta et al. (2015) observe only 2.825. This difference might be due to the low maturity of our firms, since implementing all corporate governance devices in accordance with the DCGK might take some time.

Unlike with Lopatta et al. (2015) (0.7 percent), our client importance from total sales is of substance, with a mean value of 4.1 percent. However, a median value of 0.3 percent, skewness of 3.699 and kurtosis of 17.409 imply that few extreme observations with pronounced economic bonding drive the mean, while the magnitude of the observations is located beneath the mean. We regard this as a univariate sign for distinct situations of high economic dependence that should be analyzed separately. Client importance from

[110] Untabulated tests show that $DACC_{it}$ and REM_{it} are not statistically different from zero at 0.01 levels. Mean value of 0.503 for $REMvsAEM_{it}$ also shows no systematic trade-off for the total sample.

non-audit services is 20.4 percent on average (against 31.7 percent in Lopatta et al. 2015), while an interquartile range of 30.5 percent also shows considerable variation.

Table 25 – Descriptive statistics of relevant variables

Panel: Full sample	**N**	**Mean**	Q_{50}	**Std. Dev.**	Q_{25}	Q_{75}
Dependent variables						
$DACC\_abs_{it}$	1,052	0.068	0.045	0.077	0.020	0.088
$DACC_{it}$	1,052	0.000	0.000	0.102	-0.041	0.047
REM_{it}	1,052	-0.000	-0.153	2.141	-1.226	1.109
$REMvsAEM_{it}$	1,052	0.502	0.500	0.207	0.345	0.643
Variables of interest						
CIS_{it}	1,052	0.041	0.003	0.101	0.001	0.020
$CINAS_{it}$	1,052	0.209	0.175	0.187	0.036	0.341
Control variables						
$Size_{it-1}$	1,052	10.829	10.739	1.523	9.793	11.783
MB_{it-1}	1,052	1.746	1.167	2.329	0.753	1.874
$LOSS_{it}$	1,052	0.290	0.000	0.454	0.000	1.000
$Litigation_{it}$	1,052	0.313	0.000	0.464	0.000	1.000
ACC_{it-1}	1,052	-0.052	-0.039	0.145	-0.091	0.006
CFO_{it}	1,052	0.046	0.064	0.165	0.008	0.114
$Leverage_{it-1}$	1,052	0.533	0.526	0.315	0.339	0.673
ΔTA_{it}	1,052	0.055	0.019	0.282	-0.055	0.111
$\Delta EARN_{it}$	1,052	0.019	0.002	0.184	-0.027	0.037
ROA_{it}	1,052	0.006	0.030	0.179	-0.014	0.070
$FirmAge_{it}$	1,052	2.449	2.565	0.507	2.398	2.708
$PartChange_{it}$	1,052	0.412	0.000	0.492	0.000	1.000
$AudChange_{it}$	1,052	0.134	0.000	0.341	0.000	0.000
$AudOpClean_{it}$	1,052	0.969	1.000	0.174	1.000	1.000
$BoardSize_{it}$	1,052	7.071	6.000	2.942	5.000	8.000
$BoardInd_{it}$	1,052	0.652	0.600	0.117	0.600	0.750
$VioDCGK_{it}$	1,052	6.516	6.000	4.011	4.000	9.000

All variables are defined in Appendix C.

The correlation among the most relevant variables are shown in **Table 26**. The highest correlations between independent variables is the correlation between $SIZE$ and $BoardSize$ (0.506) and ROA and $Loss$ (-0.608).[111]

[111] These results are not surprising. We conclude that there are no serious problems concerning high correlations that could majorly distort the results of our multivariate analyses. We check for the multicollinearity of our variables in all regressions.

Table 26 – Correlations among most relevant variables

			1	2	3	4	5	6	7	8	9
1.	$DACC\_abs_{it}$		1.000								
2.	REM_{it}		-0.048	1.000							
3.	$REMvsAEM_{it}$		-0.468***	-0.636***	1.000						
4.	CIS_{it}		-0.039	-0.047	0.078**	1.000					
5.	$CINAS_{it}$		0.045	-0.017	-0.024	-0.005	1.000				
6.	$Size_{it-1}$		-0.126***	0.050	0.006	-0.121***	0.240***	1.000			
7.	MB_{it-1}		0.016	-0.000	-0.040	-0.011	0.004	0.119***	1.000		
8.	$LOSS_{it}$		0.103***	-0.156***	0.063**	0.026	-0.051*	-0.288***	0.005	1.000	
9.	$Leverage_{it-1}$		-0.034	-0.095***	0.126***	0.140***	-0.007	-0.195***	-0.095***	0.171***	1.000
10.	ΔTA_{it}		-0.030	0.039	-0.041	0.006	0.131***	0.137***	0.123***	-0.237***	-0.018
11.	ROA_{it}		-0.041	0.181***	-0.134***	0.016	-0.023	0.230***	-0.068**	-0.608***	-0.211***
12.	$FirmAge_{it}$	*Ml*	-0.018	0.021	-0.007	-0.011	-0.049	-0.019	-0.010	-0.065**	0.061**
13.	$PartChange_{it}$		-0.023	-0.006	-0.005	-0.003	-0.056*	0.029	0.008	0.019	-0.015
14.	$AudChange_{it}$		0.011	-0.016	0.027	-0.019	-0.133***	-0.037	-0.040	0.050	-0.007
15.	$AudOpClean_{it}$		-0.074**	0.021	0.009	-0.022	-0.037	0.021	0.012	-0.017	-0.012
16.	$BoardSize_{it}$		-0.094***	0.065	-0.010	-0.122***	0.190***	0.506***	-0.059*	-0.138***	-0.071**
17.	$BoardInd_{it}$		-0.060*	-0.052*	0.074**	-0.020	0.025	-0.090***	-0.067**	-0.005	0.209***
18.	$VioDCGK_{it}$		0.036	-0.073**	0.043	0.046	-0.010	-0.181***	0.057*	0.069**	0.109***

Correlations among most relevant variables (continued)

		10	11	12	13	14	15	16	17	18
10.	ΔTA_{it}	1.000								
11.	ROA_{it}	0.161***	1.000							
12.	$FirmAge_{it}$	-0.075**	0.028	1.000						
13.	$PartChange_{it}$	-0.034	0.003	0.002	1.000					
14.	$AudChange_{it}$	-0.118***	-0.021	-0.005	0.425***	1.000				
15.	$AudOpClean_{it}$	0.067**	0.050	-0.032	-0.027	0.012	1.000			
16.	$BoardSize_{it}$	0.028	-0.008	0.055*	0.136	-0.013	0.020	1.000		
17.	$BoardInd_{it}$	-0.045	-0.062**	0.346***	-0.065**	-0.035	0.017	0.243***	1.000	
18.	$VioDCGK_{it}$	0.010	-0.043	-0.010	0.046	0.065**	0.040	-0.069**	0.096***	1.000

*, **, *** Indicate statistical significance at the 0.10, 0.05 and 0.01 levels, respectively, based on two-tailed tests. Variables are defined in Appendix C.

Table 27 contains descriptive statistics by audit firm size. We include this analysis to get a better impression of whether size differences within our sample of small and midsized auditors might be driving any of our results.[112] Panel A highlights that smaller audit firms are significantly higher dependent on their clients with a mean (median) value of client importance of 8.5 (2.7) percent for the smaller sized audit firms compared to 0.2 (0.1) percent for the larger sized audit firms, which is in line with the considerations by DeAngelo (1981b).[113] However, there is no significant difference in our EM proxies, which either indicates that there is no relation or a relation that cannot be captured by univariate linear testing. Interestingly, we find no significant difference for our second client importance proxy (non-audit service importance). In particular, a separate multivariate analysis of both proxies seems to be important. What is more, panel B shows that firms audited by a smaller audit firm tend to significantly trade-off AEM through REM in comparison to their counterparts audited by larger audit firms.

We split the sample once more using the respective economic bonding thresholds discussed earlier. **What is** more, we find that our two client importance proxies only once (>5% CIS in two consecutive years) show mutual significant differences. E.g., for the >10% threshold, a CIS_{it} of 26.9 percent and a $CINAS_{it}$ of 21.4 percent for the threshold firms face 1.6 percent CIS_{it} respectively 20.8 percent $CINAS_{it}$ for the remaining firms. This complements our view that both proxies measure diverse dimensions of economic bonding, which is why separate analyses are important.

[112] Size categorization (1-4) follows EU SME definition from May 2003. See Appendix for a detailed definition.
[113] The difference even increases when we put together the two groups or size 3 & 4 as the larger firms. Then, mean (median) client importance for the smaller firms is 14.1 (8.5) percent in comparison to 0.4 (0.2) percent for the larger firms.

Table 28 illustrates the statistics for each of the subsamples. We find for the CIS thresholds that firms with client importance above 5%, 10% and 15% in two consecutive years on average have comparable levels of EM. Only for the 5% threshold, univariate tests show a significant higher use of REM compared to the remaining sample. In line with this result, $REMvsAEM_{it}$ indicates that firms above the economic bonding barrier show more REM than AEM, while it is the other way round for the remaining sample. Altogether, our univariate results seem to contradict the argument that firms with a highly dependent auditor have lower audit quality and therefore lower financial reporting quality, indicating that auditors' increased reputation and litigation risk compensates the increase in dependence.

Table 27 – Descriptive statistics by audit firm size

Panel A: Full sample split by size 4

	***AudSize* 4 (1)**			***AudSize* 1-3 (2)**			**Difference (1) – (2): p-value**	
	N	Mean	Q_{50}	N	Mean	Q_{50}	Wilcoxon Test	t-test
Dependent Variables								
$DACC\_abs_{it}$	553	0.071	0.046	499	0.064	0.044	0.372	0.109
REM_{it}	553	0.011	-0.063	499	-0.012	-0.205	0.571	0.860
$REMvsAEM_{it}$	553	0.494	0.500	499	0.510	0.529	0.142	0.201
Variables of Interest								
CIS_{it}	553	0.002	0.001	499	0.085	0.024	<0.001*	<0.001*
$CISNAS_{it}$	553	0.214	0.168	499	0.203	0.179	0.773	0.349
$NumberClients_{it}$	553	23	17	499	7	3	<0.001*	<0.001*

Panel B: Full sample split by size 3-4

	***AudSize* 3-4 (1)**			***AudSize* 1-2 (2)**			**Difference (1) – (2): p-value**	
	N	Mean	Q_{50}	N	Mean	Q_{50}	Wilcoxon Test	t-test
Dependent Variables								
$DACC\_abs_{it}$	765	0.069	0.046	287	0.063	0.042	0.281	0.254
REM_{it}	765	0.012	-0.084	287	-0.032	-0.252	0.286	0.768
$REMvsAEM_{it}$	765	0.495	0.500	287	0.519	0.533	0.066*	0.092*
Variables of Interest								
CIS_{it}	765	0.004	0.002	287	0.141	0.085	<0.001*	<0.001*
$CISNAS_{it}$	765	0.212	0.175	287	0.200	0.173	0.635	0.341
$NumberClients_{it}$	765	20	16	287	2	1	<0.001*	<0.001*

To split firm-year observations by the size of the audit firm, we use a size system that follows EU SME definitions from May 2003. Therefore, we define audit firms with sales <=2 m. EUR as midget audit firms (size 1) and audit firms with sales >2 m. EUR and <=10 m. EUR as small audit firms (size 2). We define audit firms with sales >10 m. EUR and <=50 m. EUR as medium audit firms (size 3) and audit firms with sales > 50 m. EUR as large audit firms (size 4). We evaluate significances of means and medians based on t-tests and Wilcoxon tests, respectively. We calculate p-values and Z-statistics based on two-tailed tests. * indicates statistical significance at conventional levels. Variables are defined in Appendix C.

What is more, we find that our two client importance proxies only once (>5% CIS in two consecutive years) show mutual significant differences. E.g., for the >10% threshold, a CIS_{it} of 26.9 percent and a $CINAS_{it}$ of 21.4 percent for the threshold firms face 1.6 percent CIS_{it} respectively 20.8 percent $CINAS_{it}$ for the remaining firms. This complements our view that both proxies measure diverse dimensions of economic bonding, which is why separate analyses are important.

Table 28 – Descriptive statistics by CI limits for the full sample

	>5% CIS in two consecutive years (1)			**Remaining sample (2)**			**Difference (1) – (2): p-value**	
	N	Mean	Q_{50}	N	Mean	Q_{50}	Wilcoxon Test	t-test
Dependent Variables								
$DACC\_abs_{it}$	171	0.062	0.043	881	0.069	0.045	0.487	0.252
REM_{it}	171	-0.182	-0.431	881	0.035	-0.087	0.046*	0.224
$REMvsAEM_{it}$	171	0.523	0.556	881	0.498	0.500	0.092*	0.144
Variables of Interest								
CIS_{it}	171	0.198	0.139	881	0.011	0.002	<0.001*	<0.001*
$CISNAS_{it}$	171	0.233	0.230	881	0.204	0.167	0.028*	0.063*
	>10% CIS in two consecutive years (1)			**Remaining sample (2)**			**Difference (1) – (2): p-value**	
	N	Mean	Q_{50}	N	Mean	Q_{50}	Wilcoxon Test	t-test
Dependent Variables								
$DACC\_abs_{it}$	103	0.064	0.047	949	0.068	0.045	0.990	0.588
REM_{it}	103	-0.249	-0.367	949	0.027	-0.124	0.127	0.213
$REMvsAEM_{it}$	103	0.519	0.538	949	0.500	0.500	0.342	0.381
Variables of Interest								
CIS_{it}	103	0.269	0.203	949	0.016	0.002	<0.001*	<0.001*
$CISNAS_{it}$	103	0.214	0.220	949	0.208	0.170	0.701	0.781
	>15% CIS in two consecutive years (1)			**Remaining sample (2)**			**Difference (1) – (2): p-value**	
	N	Mean	Q_{50}	N	Mean	Q_{50}	Wilcoxon Test	t-test
Dependent Variables								
$DACC\_abs_{it}$	70	0.066	0.046	982	0.068	0.045	>0.999	0.882
REM_{it}	70	-0.290	-0.326	982	0.021	-0.135	0.189	0.241
$REMvsAEM_{it}$	70	0.524	0.542	982	0.500	0.500	0.331	0.357
Variables of Interest								
CIS_{it}	70	0.330	0.273	982	0.021	0.003	<0.001*	<0.001*
$CISNAS_{it}$	70	0.202	0.213	982	0.209	0.173	0.637	0.76
	>40% CINAS in two consecutive years (1)			**Remaining sample (2)**			**Difference (1) – (2): p-value**	
	N	Mean	Q_{50}	N	Mean	Q_{50}	Wilcoxon Test	t-test
Dependent Variables								
$DACC\_abs_{it}$	94	0.076	0.054	958	0.067	0.044	0.163	0.249
REM_{it}	94	0.020	-0.052	958	-0.002	-0.161	0.605	0.92
$REMvsAEM_{it}$	94	0.473	0.500	958	0.505	0.500	0.161	0.164
Variables of Interest								
CIS_{it}	94	0.055	0.005	958	0.040	0.003	0.094*	0.164
$CISNAS_{it}$	94	0.534	0.513	958	0.177	0.152	<0.001*	<0.001*
	>50% CINAS in two consecutive years (1)			**Remaining sample (2)**			**Difference (1) – (2): p-value**	
	N	Mean	Q_{50}	N	Mean	Q_{50}	Wilcoxon Test	t-test
Dependent Variables								
$DACC\_abs_{it}$	38	0.088	0.050	1,014	0.067	0.045	0.631	0.104
REM_{it}	38	0.020	-0.083	1,014	-0.001	-0.157	0.486	0.952
$REMvsAEM_{it}$	38	0.479	0.464	1,014	0.503	0.500	0.389	0.491
Variables of Interest								
CIS_{it}	38	0.017	0.005	1,014	0.042	0.003	0.975	0.128
$CISNAS_{it}$	38	0.625	0.615	1,014	0.193	0.167	<0.001*	<0.001*

A firm is defined as a dependent firm when the client importance concerning the respective proxy (CIS or $CINAS$) lies above the respective threshold in two consecutive years, whereas all other observations are attributed to the remaining sample. We evaluate significances of means and medians based on t-tests and Wilcoxon tests, respectively. We calculate p-values and Z-statistics based on two-tailed tests. * indicates statistical significance at conventional levels. Variables are defined in Appendix C.

3.6 Multivariate analysis

We conduct multivariate regressions on the influence of client importance from sales (CIS_{it}) using **Equation 24-Equation 25** and report the results in **Table 29**, **Table 30** and **Table 31**. All regressions are statistically significant at the 0.01 levels and explanatory power (R^2) ranges from 5-16 percent using cross-sectional EM proxies. These numbers are on average slightly higher than with Sharma et al. (2011), though slightly lower than with Svanström (2013). Results in **Table 29** show the multiple regressions of AEM, e.g., $DACC\_abs_{it}$, on CIS. Columns (1)-(3) use a DEP_{it} indicator with client importance above 5% in one, two or three consecutive years. Coefficients on CIS_{it} are negative and significant, indicating that firms with higher dependence tend to use lower discretionary accruals. Initially, this opposes the hypothesized link that firms with higher dependence can use this dependence in financial reporting negotiations to pursue more AEM. Furthermore, this works in favor of our argument for higher reputation and litigation risk for publicly listed clients as fees increase, which seems to overcompensate increased economic dependence. The suspected high-dependence firms with above 5% total sales importance show mixed results for the coefficients on DEP_{it} and the interaction variable ($CIS_{it} * DEP_{it}$). We find a significant negative coefficient on DEP_{it}, while the coefficient on $CIS_{it} * DEP_{it}$ is significant and positive for one and three consecutive years. We regard this as a weak sign for an increase of the economic dependence effect as the dependence gets pronounced, e.g., exceeds the regulatory thresholds.

Table 29 – Multiple regressions of AEM on CIS

Dependent variable: absolute positively scaled value of discretionary accruals: $DACC\_abs_{it}$									
	>5 % *CIS*			**>10 % *CIS***			**>15 % *CIS***		
	(1) one year	(2) two years	(3) three years	(4) one year	(5) two years	(6) three years	(7) one year	(8) two years	(9) three years
CIS_{it}	-0.716 (3.83)***	-0.105 (3.24)***	-0.095 (3.32)***	-0.219 (2.22)**	-0.120 (3.34)***	-0.119 (3.93)***	-0.161 (2.62)***	-0.115 (3.94)***	-0.106 (4.78)***
DEP_{it}	-0.012 (1.76)*	-0.013 (1.78)*	-0.018 (2.19)**	-0.016 (2.26)**	-0.018 (1.66)*	-0.021 (1.86)*	-0.015 (1.10)	-0.016 (0.71)	-0.029 (1.30)
$CIS_{it} * DEP_{it}$	0.700 (3.77)***	0.108 (1.59)	0.117 (1.78)*	0.221 (2.10)**	0.143 (1.89)*	0.161 (2.69)***	0.161 (2.45)**	0.132 (1.58)	0.162 (2.09)**
REM_{it}	-0.001 (0.90)	-0.001 (0.90)	-0.001 (0.90)	-0.001 (0.91)	-0.001 (0.88)	-0.001 (0.78)	-0.001 (0.91)	-0.001 (0.87)	-0.001 (0.77)
$Size_{it-1}$	-0.005 (2.48)**	-0.005 (2.38)**	-0.005 (2.36)**	-0.005 (2.37)**	-0.005 (2.32)**	-0.005 (2.33)**	-0.005 (2.48)**	-0.005 (2.37)**	-0.005 (2.44)**
MB_{it-1}	0.000 (0.34)	0.000 (0.34)	0.000 (0.33)	0.000 (0.34)	0.000 (0.36)	0.000 (0.31)	0.000 (0.36)	0.000 (0.38)	0.000 (0.35)
$LOSS_{it}$	0.016 (1.83)*	0.016 (1.82)*	0.016 (1.86)*	0.016 (1.88)*	0.016 (1.83)*	0.016 (1.86)*	0.016 (1.87)*	0.016 (1.79)*	0.016 (1.84)*
$Litigation_{it}$	-0.001 (0.25)	-0.001 (0.18)	-0.001 (0.20)	-0.001 (0.21)	-0.001 (0.23)	-0.001 (0.24)	-0.001 (0.22)	-0.001 (0.23)	-0.002 (0.26)
ACC_{it-1}	-0.025 (0.67)	-0.024 (0.67)	-0.024 (0.66)	-0.024 (0.66)	-0.026 (0.71)	-0.025 (0.71)	-0.023 (0.65)	-0.026 (0.73)	-0.025 (0.71)
CFO_{it}	-0.019 (0.39)	-0.017 (0.35)	-0.017 (0.34)	-0.017 (0.35)	-0.017 (0.36)	-0.023 (0.46)	-0.017 (0.34)	-0.018 (0.36)	-0.022 (0.45)
$Leverage_{it-1}$	-0.015 (1.03)	-0.015 (1.01)	-0.015 (1.02)	-0.015 (1.03)	-0.014 (0.90)	-0.013 (0.83)	-0.015 (1.02)	-0.014 (0.89)	-0.013 (0.88)
ΔTA_{it}	-0.000 (0.02)	0.001 (0.04)	0.001 (0.05)	0.000 (0.03)	0.000 (0.03)	0.000 (0.00)	0.001 (0.05)	0.000 (0.03)	0.000 (0.02)
$\Delta EARN_{it}$	0.003 (0.15)	0.004 (0.22)	0.004 (0.26)	0.004 (0.23)	0.003 (0.17)	0.003 (0.19)	0.004 (0.25)	0.002 (0.16)	0.003 (0.19)
ROA_{it}	0.031 (0.70)	0.028 (0.63)	0.027 (0.61)	0.028 (0.65)	0.029 (0.66)	0.033 (0.75)	0.029 (0.65)	0.029 (0.67)	0.033 (0.75)
$FirmAge_{it}$	0.001 (0.35)	0.002 (0.42)	0.002 (0.47)	0.001 (0.39)	0.001 (0.41)	0.002 (0.46)	0.001 (0.41)	0.001 (0.36)	0.001 (0.41)
$PartChange_{it}$	-0.005 (1.36)	-0.005 (1.39)	-0.005 (1.42)	-0.005 (1.41)	-0.005 (1.44)	-0.006 (1.54)	-0.005 (1.48)	-0.005 (1.44)	-0.006 (1.52)
$AudChange_{it}$	0.001 (0.12)	0.003 (0.36)	0.003 (0.31)	0.002 (0.21)	0.004 (0.43)	0.004 (0.44)	0.002 (0.23)	0.004 (0.42)	0.004 (0.39)
$AudOpClean_{it}$	-0.035 (1.44)	-0.034 (1.40)	-0.035 (1.43)	-0.035 (1.43)	-0.034 (1.40)	-0.034 (1.39)	-0.034 (1.41)	-0.034 (1.38)	-0.035 (1.43)
$BoardSize_{it}$	-0.001 (0.89)	-0.001 (0.89)	-0.001 (0.93)	-0.001 (0.97)	-0.001 (0.90)	-0.001 (0.95)	-0.001 (0.97)	-0.001 (0.91)	-0.001 (0.93)
$BoardInd_{it}$	-0.036 (1.23)	-0.034 (1.15)	-0.034 (1.14)	-0.035 (1.17)	-0.033 (1.12)	-0.033 (1.11)	-0.034 (1.15)	-0.033 (1.11)	-0.033 (1.11)
$VioDCGK_{it}$	0.000 (0.70)	0.000 (0.70)	0.000 (0.63)	0.000 (0.61)	0.000 (0.74)	0.000 (0.75)	0.000 (0.61)	0.000 (0.75)	0.000 (0.74)
$Intercept$	0.188 (3.74)***	0.181 (3.57)***	0.181 (3.60)***	0.185 (3.63)***	0.180 (3.41)***	0.179 (3.42)***	0.185 (3.59)***	0.180 (3.46)***	0.182 (3.50)***
$Industry\ FE$	Included	Included	Included	Included	Included	Included	Included	Included	Included
$Year\ FE$	Included	Included	Included	Included	Included	Included	Included	Included	Included
R^2	0.05	0.05	0.05	0.05	0.05	0.05	0.05	0.05	0.05
F	2.82	2.76	2.95	2.63	2.91	3.23	2.62	2.86	2.96
n	1,052	1,052	1,052	1,052	1,052	1,052	1,052	1,052	1,052

*, **, *** Indicate statistical significance at the 0.10, 0.05, and 0.01 levels, respectively, based on two-tailed tests. All test statistics are calculated using OLS with two-way clustering at the firm and time level. We additionally include appropriate fixed effects. Mean VIF for all regressions are considerably below 5. $DACC\_abs$ is comprehended as the residual from cross-sectional industry-year specific first-stage regressions of Equation 12. CIS is client importance, calculated as ratio of total client sales to total sales of the audit firm. All other variables are defined in Appendix C.

Still, whenever we find a significant coefficient on DEP_{it}, the total effects remains negative, e.g., the economic dependence effect is still offset by an increased effect of reputation and litigation risk, only by a decreased margin. This can be seen from simple calculation of $\frac{|\beta_2|}{|\beta_3-\beta_1|}$, indicating a turning point for dependence levels outside of the distribution.[114]

Taking a closer look at the 10% (column (4)-(6)) thresholds, we find similar results. What is important and in line with considerations of an increasing economic dependence as the time period increases, we find for all regressions that coefficient β_1 gradually decreases, while coefficient β_3 increases. We propose that as the number EM incidents increases, the tendency of the auditor to rule these out decreases as the dependence on the client increases.

When we consider the results for the 15% threshold in columns (7)-(9), results change considerably. For firms in this threshold, there in an overcompensating dependence effects ($\beta_3 \geq \beta_1$), especially for a time-period of three consecutive years. We conclude that exceeding both the high dependence level (e.g., 15% CIS) and the long time-period could be a drawback for audit quality and might open up opportunities for the client to pursue AEM.[115]

Results on REM_{it} are reported in **Table 30**. Being negatively scaled, higher auditor dependence from clients would go hand in hand with more REM when coefficients are negative. However, we do not find significant results for an influence of client importance on the use of REM. Following Greiner et al. (2017), we note that this result could be due to the fact that audit characteristics only influence the extreme parts of the distribution of REM, which is why we again focus on this part of EM in our additional analysis in section 3.7.

[114] E.g., for >5 % CIS for three consecutive years (column (3)), the total effects turns positive when dependence exceeds a level of 81.82 percent, while maximum CIS is 57.57 percent.
[115] However, we not that only few observations with a dependence threshold above 15% in three consecutive years exist. Therefore, we consecutively test the robustness of our results using pooled AEM proxies and in section 3.7 for extreme observations of AEM.

Table 30 – Multiple regressions of REM on CIS

Dependent variable: comprehensive negatively scaled standardized *REM* proxy: $REM_{it} = CFO\_abn_{it} - PROD\_abn_{it} + DISEXP\_abn_{it}$									
	>5 % *CIS*			**>10 % *CIS***			**>15 % *CIS***		
	(1) one year	(2) two years	(3) three years	(4) one year	(5) two years	(6) three years	(7) one year	(8) two years	(9) three years
CIS_{it}	-0.889	-1.338	-1.278	-1.219	-0.550	0.536	-2.114	-0.480	0.565
	(0.11)	(1.31)	(1.27)	(0.25)	(0.62)	(0.36)	(0.62)	(0.39)	(0.39)
DEP_{it}	-0.303	-0.264	-0.189	-0.535	-0.272	0.104	-0.412	-0.255	0.100
	(0.75)	(0.58)	(0.36)	(1.08)	(0.55)	(0.20)	(1.05)	(0.63)	(0.20)
$CIS_{it} * DEP_{it}$	1.348	1.854	1.640	2.237	0.927	-1.508	2.805	0.749	-1.594
	(0.18)	(0.85)	(0.73)	(0.40)	(0.48)	(0.61)	(0.74)	(0.50)	(0.70)
$DACC\_abs_{it}$	-0.849	-0.870	-0.872	-0.868	-0.845	-0.761	-0.864	-0.832	-0.761
	(0.86)	(0.86)	(0.86)	(0.87)	(0.85)	(0.75)	(0.87)	(0.84)	(0.75)
$Size_{it-1}$	-0.052	-0.051	-0.050	-0.060	-0.055	-0.052	-0.055	-0.053	-0.052
	(0.66)	(0.64)	(0.64)	(0.74)	(0.67)	(0.65)	(0.70)	(0.68)	(0.67)
MB_{it-1}	0.023	0.023	0.023	0.026	0.025	0.025	0.025	0.024	0.024
	(0.45)	(0.44)	(0.44)	(0.48)	(0.46)	(0.47)	(0.47)	(0.46)	(0.47)
$LOSS_{it}$	-0.341	-0.346	-0.341	-0.342	-0.340	-0.334	-0.341	-0.340	-0.335
	(1.98)**	(2.01)**	(1.95)*	(1.96)*	(1.95)*	(1.90)*	(1.97)**	(1.92)*	(1.89)*
$Litigation_{it}$	0.393	0.392	0.393	0.383	0.394	0.407	0.387	0.396	0.409
	(1.61)	(1.61)	(1.61)	(1.56)	(1.60)	(1.64)	(1.58)	(1.61)	(1.66)*
ACC_{it-1}	0.019	0.007	0.008	0.031	0.017	0.051	0.017	0.018	0.056
	(0.05)	(0.02)	(0.02)	(0.10)	(0.06)	(0.17)	(0.05)	(0.06)	(0.19)
CFO_{it}	6.280	6.271	6.261	6.276	6.259	6.295	6.268	6.252	6.301
	(6.50)***	(6.47)***	(6.44)***	(6.57)***	(6.49)***	(6.42)***	(6.48)***	(6.46)***	(6.41)***
$Leverage_{it-1}$	0.279	0.279	0.280	0.270	0.283	0.268	0.273	0.283	0.266
	(0.62)	(0.63)	(0.63)	(0.61)	(0.65)	(0.62)	(0.61)	(0.65)	(0.62)
ΔTA_{it}	0.352	0.351	0.354	0.356	0.356	0.367	0.354	0.358	0.368
	(1.66)*	(1.68)*	(1.70)*	(1.63)	(1.70)*	(1.75)*	(1.68)*	(1.74)*	(1.79)*
$\Delta EARN_{it}$	0.475	0.466	0.467	0.473	0.472	0.493	0.468	0.473	0.498
	(1.06)	(1.03)	(1.04)	(1.13)	(1.13)	(1.17)	(1.08)	(1.13)	(1.18)
ROA_{it}	-2.348	-2.343	-2.327	-2.334	-2.307	-2.322	-2.326	-2.302	-2.326
	(2.07)**	(2.07)**	(2.04)**	(2.08)**	(2.03)**	(2.08)**	(2.05)**	(2.01)**	(2.08)**
$FirmAge_{it}$	-0.018	-0.017	-0.015	-0.020	-0.021	-0.024	-0.020	-0.021	-0.019
	(0.10)	(0.09)	(0.08)	(0.11)	(0.12)	(0.13)	(0.11)	(0.12)	(0.11)
$PartChange_{it}$	0.024	0.025	0.021	0.024	0.024	0.023	0.025	0.023	0.022
	(0.18)	(0.19)	(0.16)	(0.19)	(0.19)	(0.18)	(0.19)	(0.18)	(0.18)
$AudChange_{it}$	-0.073	-0.055	-0.050	-0.080	-0.063	-0.074	-0.076	-0.059	-0.072
	(0.32)	(0.23)	(0.20)	(0.34)	(0.26)	(0.29)	(0.33)	(0.26)	(0.31)
$AudOpClean_{it}$	0.181	0.189	0.197	0.182	0.194	0.207	0.178	0.189	0.202
	(0.47)	(0.49)	(0.52)	(0.46)	(0.50)	(0.54)	(0.46)	(0.50)	(0.54)
$BoardSize_{it}$	0.035	0.036	0.036	0.036	0.037	0.038	0.036	0.037	0.038
	(0.94)	(0.95)	(0.96)	(0.98)	(0.99)	(1.02)	(0.97)	(1.00)	(1.03)
$BoardInd_{it}$	-0.858	-0.852	-0.858	-0.839	-0.846	-0.882	-0.853	-0.859	-0.895
	(0.85)	(0.84)	(0.85)	(0.84)	(0.84)	(0.87)	(0.85)	(0.85)	(0.88)
$VioDCGK_{it}$	-0.016	-0.015	-0.015	-0.016	-0.015	-0.014	-0.016	-0.014	-0.014
	(0.85)	(0.77)	(0.76)	(0.84)	(0.75)	(0.74)	(0.82)	(0.75)	(0.74)
$Intercept$	0.427	0.386	0.360	0.506	0.400	0.362	0.466	0.388	0.352
	(0.41)	(0.38)	(0.35)	(0.48)	(0.38)	(0.34)	(0.46)	(0.39)	(0.35)
Industry FE	Included	Included	Included	Included	Included	Included	Included	Included	Included
Year FE	Included	Included	Included	Included	Included	Included	Included	Included	Included
R^2	0.16	0.16	0.16	0.16	0.16	0.16	0.16	0.16	0.16
F	6.13	6.42	6.43	6.14	6.08	6.31	6.06	6.03	6.28
n	1,052	1,052	1,052	1,052	1,052	1,052	1,052	1,052	1,052

*, **, *** Indicate statistical significance at the 0.10, 0.05, and 0.01 levels, respectively, based on two-tailed tests. All test statistics are calculated using OLS with two-way clustering at the firm and time level. We additionally include appropriate fixed effects. Mean VIF for all regressions are considerably below 5. *REM* is comprehended as the standardized value of the residuals from cross-sectional industry-year specific first-stage regressions of Equation 13,Equation 14 and Equation 17. *CIS* is client importance, calculated as ratio of total client sales to total sales of the audit firm. All other variables are defined in Appendix C.

When we take a closer look at the trade-off relation proxy ($REMvsAEM_{it}$) in **Table 31**, we find a positive and significant effect of higher client importance, indicating that higher dependence leads to a trade-off from AEM to real activities. These results are only consistent for cases in which we study the respective threshold for two consecutive years. Here we find significant (at least 0.05 levels) results for a trade-off from AEM to REM, which is decreasing with an increasing dependence threshold, indicating the already discussed increase in dependence effects with increasing dependence threshold. Furthermore, for 10% and 15% thresholds, we also find a significant negative coefficients for our interaction term (columns (5) & (8)), indicating a trade-off back to AEM when the threshold is met.

Altogether, these results work against the understanding that a less independent auditor will yield opportunities for the client to purse AEM. In contrast, increasing economic dependence seems to enhance audit quality, e.g., the higher dependence as an input factor causes higher audit effort due to the pronounced reputation and litigation risk effects, causing the client to use lower AEM. Consequently, clients could take into consideration the trade-off between AEM and REM (e.g., Zang 2012) and use costly REM to compensate decreased opportunities of AEM.

Table 31 – Multiple regressions of the trade-off proxy REMvsAEM on CIS

Dependent variable: decile-based trade-off proxy between *REM* and *AEM*: $REMvsAEM_{it}$									
	>5 % *CIS*			**>10 % *CIS***			**>15 % *CIS***		
	(1) one year	(2) two years	(3) three years	(4) one year	(5) two years	(6) three years	(7) one year	(8) two years	(9) three years
CIS_{it}	1.579	0.372	0.326	0.591	0.357	0.260	0.362	0.288	0.218
	(2.23)**	(2.23)**	(2.26)**	(1.26)	(2.80)***	(1.59)	(1.28)	(2.57)**	(1.52)
DEP_{it}	0.030	0.031	0.044	0.039	0.034	0.036	0.043	0.046	0.048
	(1.08)	(1.05)	(1.40)	(1.83)*	(1.55)	(1.14)	(3.91)***	(2.69)***	(1.34)
$CIS_{it} * DEP_{it}$	-1.540	-0.391	-0.393	-0.587	-0.394	-0.299	-0.367	-0.341	-0.272
	(2.21)**	(1.55)	(1.64)	(1.26)	(2.44)**	(1.22)	(1.46)	(2.20)**	(1.11)
$Size_{it-1}$	0.012	0.012	0.011	0.012	0.012	0.012	0.012	0.012	0.012
	(2.11)**	(2.03)**	(2.00)**	(2.03)**	(2.03)**	(2.06)**	(2.17)**	(2.13)**	(2.16)**
MB_{it-1}	-0.005	-0.005	-0.005	-0.005	-0.005	-0.005	-0.005	-0.005	-0.005
	(1.53)	(1.54)	(1.49)	(1.49)	(1.50)	(1.45)	(1.59)	(1.60)	(1.56)
$LOSS_{it}$	-0.013	-0.012	-0.013	-0.013	-0.012	-0.013	-0.013	-0.013	-0.013
	(0.68)	(0.62)	(0.65)	(0.72)	(0.67)	(0.71)	(0.73)	(0.68)	(0.71)
$Litigation_{it}$	-0.019	-0.019	-0.019	-0.019	-0.019	-0.019	-0.019	-0.018	-0.019
	(0.97)	(0.97)	(0.96)	(0.99)	(0.96)	(0.98)	(0.99)	(0.95)	(0.98)
ACC_{it-1}	-0.018	-0.017	-0.018	-0.019	-0.013	-0.017	-0.021	-0.014	-0.018
	(0.37)	(0.37)	(0.40)	(0.40)	(0.28)	(0.39)	(0.46)	(0.34)	(0.43)
CFO_{it}	-0.391	-0.393	-0.395	-0.394	-0.394	-0.385	-0.395	-0.394	-0.386
	(5.73)***	(5.73)***	(5.77)***	(5.66)***	(5.85)***	(5.57)***	(5.76)***	(5.88)***	(5.78)***
$Leverage_{it-1}$	0.027	0.027	0.027	0.027	0.023	0.022	0.027	0.024	0.024
	(0.89)	(0.90)	(0.92)	(0.92)	(0.78)	(0.76)	(0.93)	(0.83)	(0.80)
ΔTA_{it}	-0.037	-0.039	-0.039	-0.038	-0.038	-0.038	-0.039	-0.038	-0.039
	(1.94)*	(2.17)**	(2.20)**	(1.99)**	(2.04)**	(2.06)**	(2.13)**	(2.11)**	(2.15)**
$\Delta EARN_{it}$	-0.047	-0.047	-0.048	-0.049	-0.045	-0.048	-0.049	-0.046	-0.048
	(0.97)	(0.99)	(1.03)	(1.05)	(0.97)	(1.03)	(1.10)	(1.01)	(1.05)
ROA_{it}	0.080	0.087	0.089	0.086	0.086	0.076	0.085	0.085	0.077
	(1.02)	(1.10)	(1.13)	(1.09)	(1.08)	(0.95)	(1.09)	(1.08)	(0.99)
$FirmAge_{it}$	-0.003	-0.004	-0.005	-0.004	-0.004	-0.004	-0.004	-0.003	-0.004
	(0.24)	(0.29)	(0.33)	(0.27)	(0.28)	(0.29)	(0.27)	(0.24)	(0.27)
$PartChange_{it}$	-0.011	-0.010	-0.010	-0.010	-0.010	-0.009	-0.010	-0.010	-0.009
	(1.09)	(1.03)	(0.96)	(1.02)	(1.01)	(0.94)	(0.99)	(1.03)	(0.97)
$AudChange_{it}$	0.026	0.018	0.019	0.024	0.017	0.019	0.023	0.019	0.020
	(1.18)	(0.68)	(0.75)	(1.11)	(0.65)	(0.71)	(1.07)	(0.72)	(0.77)
$AudOpClean_{it}$	0.018	0.015	0.016	0.018	0.014	0.015	0.017	0.015	0.016
	(0.70)	(0.60)	(0.64)	(0.71)	(0.59)	(0.59)	(0.66)	(0.57)	(0.59)
$BoardSize_{it}$	-0.003	-0.003	-0.003	-0.003	-0.003	-0.003	-0.003	-0.003	-0.003
	(1.12)	(1.00)	(0.97)	(1.01)	(1.00)	(1.01)	(1.02)	(1.02)	(1.03)
$BoardInd_{it}$	0.113	0.108	0.108	0.111	0.106	0.107	0.109	0.106	0.108
	(1.50)	(1.40)	(1.38)	(1.48)	(1.39)	(1.38)	(1.42)	(1.36)	(1.36)
$VioDCGK_{it}$	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001
	(0.68)	(0.57)	(0.67)	(0.75)	(0.53)	(0.56)	(0.73)	(0.59)	(0.59)
$Intercept$	0.316	0.334	0.334	0.322	0.338	0.337	0.323	0.332	0.332
	(4.19)***	(4.44)***	(4.51)***	(4.31)***	(4.24)***	(4.20)***	(4.36)***	(4.30)***	(4.45)***
$Industry\ FE$	Included	Included	Included	Included	Included	Included	Included	Included	Included
$Year\ FE$	Included	Included	Included	Included	Included	Included	Included	Included	Included
R^2	0.09	0.09	0.09	0.09	0.09	0.09	0.09	0.09	0.09
F	5.57	5.68	5.82	5.47	5.76	4.98	5.36	5.61	5.04
n	1,052	1,052	1,052	1,052	1,052	1,052	1,052	1,052	1,052

*, **, *** Indicate statistical significance at the 0.10, 0.05, and 0.01 levels, respectively, based on two-tailed tests. All test statistics are calculated using OLS with two-way clustering at the firm and time level. We additionally include appropriate fixed effects. Mean VIF for all regressions are considerably below 5. *REMvsAEM* is comprehended as the ratio of the decile of *REM* to the sum of the deciles of *REM* and *DACC_abs*, where *REM* is the standardized value of the residuals from cross-sectional industry-year specific first-stage regressions of Equation 13, Equation 14 and Equation 17 *DACC_abs* is the residual from cross-sectional industry-year specific first-stage regressions of Equation 12. *CIS* is client importance, calculated as ratio of total client sales to total sales of the audit firm. All other variables are defined in Appendix C.

As already mentioned in section 2.2.1, results from first-stage cross-sectional regressions might lack efficiency, which is why we alternatively use pooled first-stage regressions with two-way clustering at the firm and year level. Using residuals from these regressions in **Table 32**, adjusted r-squares are considerably higher between 12-31 percent.[116] We find significant and reassuring results for AEM in columns (1)-(2), while for REM we find significant negative coefficients in columns (3)-(4), indicating a continuous increase of REM with increasing dependence of the auditor. The results for our trade-off proxy support this and show a holistic picture of an increasing audit quality when dependence increases, which is why clients seem to trade-off from AEM to REM. On the other hand, audit quality seems to decrease when thresholds are met, followed by a higher use of AEM while REM seems to remain increasing, though at a lower level, therefore decreasing the trade-off relation between REM and AEM again.

We test the influence of the relation of non-audit fees to total fees from a specific audit client ($CINAS$) on EM and report the regression results in **Table 33**. The binary variable DEP is 1 if non-audit fee ratios are larger than 40% or 50%, respectively. Again, we focus on two consecutive years as the respective time-period of dependence.[117] The results show that an increasing client importance, proxied by the non-audit fee to total fee ratio, does not significantly explain any variation in our EM proxies. Conjectured interaction effects (non-linearities) are not significant. If only, the results in column (4) for the dependence threshold and the separate influence of $CINAS$ for this group show results, which are in line with our results for our first client importance proxy. This result is in line with our null hypotheses. However, it is important to note that these regression results are opposed to the results on CIS. Given that CIS is a much more precise and particularly direct proxy for client importance, we note that studies using non-audit fee importance should be careful when it comes to drawing conclusion about client importance using this proxy. There might be diverse incentives captured by the $CINAS$ proxy, which could lead to false inferences, at least for comparable non-BigN settings.

[116] To ensure clarity of our study, we only report results for 5% and 10% for two consecutive years. Results for one and three consecutive years are comparable, while results for the 15% threshold lack number of firm-years with this feature, again reducing the efficiency of our estimations.
[117] The results for one or three consecutive years are qualitatively the same, which is why we do not report them.

Table 32 – Multiple regressions of alternative EM proxies on CIS for two consecutive years

	Dependent variable: $DACC\_abs_{it}$		Dependent variable: REM_{it}		Dependent variable: $REMvsAEM_{it}$	
	(1)	(2)	(3)	(4)	(5)	(6)
	>5 % CIS	>10 % CIS	>5 % CIS	>10 % CIS	>5 % CIS	>10 % CIS
CIS_{it}	-0.119	-0.109	-1.798	-1.780	0.196	0.212
	(8.09)***	(3.50)***	(1.72)*	(2.27)**	(2.59)***	(2.45)**
DEP_{it}	-0.020	-0.044	-0.261	-0.352	0.059	0.098
	(2.25)**	(2.77)***	(0.79)	(0.87)	(2.14)**	(2.40)**
$CIS_{it} * DEP_{it}$	0.148	0.203	2.085	2.335	-0.280	-0.400
	(4.02)***	(3.59)***	(1.32)	(1.39)	(2.80)***	(4.05)***
REM_{it}	-0.001	-0.001				
	(1.03)	(1.08)				
$DACC\_abs_{it}$			-0.451	-0.464		
			(1.02)	(1.07)		
$Size_{it-1}$	-0.011	-0.012	-0.137	-0.139	0.027	0.028
	(3.06)***	(3.02)***	(1.84)*	(1.87)*	(4.02)***	(4.15)***
MB_{it-1}	0.001	0.001	0.061	0.062	-0.009	-0.009
	(1.81)*	(1.82)*	(1.51)	(1.50)	(2.62)***	(2.62)***
$LOSS_{it}$	0.018	0.018	-0.270	-0.268	-0.011	-0.011
	(1.83)*	(1.77)*	(1.60)	(1.57)	(0.45)	(0.42)
$Litigation_{it}$	-0.007	-0.008	0.988	0.985	-0.049	-0.047
	(0.85)	(0.90)	(5.22)***	(5.17)***	(2.29)**	(2.20)**
ACC_{it-1}	-0.029	-0.030	-0.749	-0.770	0.032	0.035
	(0.55)	(0.57)	(1.47)	(1.67)*	(0.75)	(0.91)
CFO_{it}	-0.115	-0.115	7.455	7.451	-0.443	-0.441
	(0.92)	(0.93)	(13.76)***	(14.48)***	(4.37)***	(4.50)***
$Leverage_{it-1}$	0.014	0.015	-0.079	-0.063	-0.009	-0.010
	(0.94)	(0.93)	(0.17)	(0.14)	(0.26)	(0.29)
ΔTA_{it}	0.040	0.040	0.345	0.345	-0.086	-0.087
	(1.93)*	(1.98)**	(1.61)	(1.61)	(3.25)***	(3.29)***
$\Delta EARN_{it}$	0.023	0.023	0.180	0.169	-0.050	-0.047
	(0.66)	(0.67)	(1.24)	(11.59)***	(1.37)	(1.45)
ROA_{it}	0.053	0.054	-2.485	-2.464	0.148	0.143
	(0.53)	(0.55)	(2.79)***	(2.82)***	(1.61)	(1.61)
$FirmAge_{it}$	-0.004	-0.005	0.157	0.154	0.006	0.007
	(0.63)	(0.68)	(0.80)	(0.78)	(0.35)	(0.39)
$PartChange_{it}$	-0.011	-0.010	-0.043	-0.042	-0.007	-0.008
	(1.79)*	(1.80)*	(0.42)	(0.42)	(0.67)	(0.72)
$AudChange_{it}$	0.007	0.006	-0.137	-0.130	0.006	0.005
	(0.58)	(0.53)	(1.18)	(1.26)	(0.32)	(0.25)
$AudOpClean_{it}$	-0.024	-0.025	0.306	0.309	0.040	0.039
	(0.90)	(0.93)	(0.77)	(0.75)	(1.44)	(1.35)
$BoardSize_{it}$	-0.002	-0.002	-0.010	-0.010	0.001	0.001
	(1.32)	(1.27)	(0.21)	(0.21)	(0.46)	(0.44)
$BoardInd_{it}$	-0.032	-0.029	0.115	0.132	0.014	0.008
	(0.97)	(0.90)	(0.14)	(0.16)	(0.16)	(0.10)
$VioDCGK_{it}$	-0.001	-0.001	-0.012	-0.012	0.002	0.002
	(0.68)	(0.70)	(0.68)	(0.66)	(1.55)	(1.45)
$Intercept$	0.267	0.271	0.309	0.313	0.201	0.193
	(3.16)***	(3.09)***	(0.37)	(0.35)	(1.65)*	(1.52)
$Industry\ FE$	Included	Included	Included	Included	Included	Included
$Year\ FE$	Included	Included	Included	Included	Included	Included
R^2	0.12	0.13	0.31	0.31	0.14	0.14
F	5.90	5.87	16.83	17.24	8.43	8.41
n	1,052	1,052	1,052	1,052	1,052	1,052

*, **, *** Indicate statistical significance at the 0.10, 0.05, and 0.01 levels, respectively, based on two-tailed tests. All test statistics are calculated using OLS with two-way clustering at the firm and time level. We additionally include appropriate fixed effects. Mean VIF for all regressions are considerably below 5. $DACC\_abs$ (REM) is comprehended as the residual from pooled first-stage regressions of Equation 12 (Equation 13, Equation 14 and Equation 17) with two-way clustering. $REMvsAEM$ is comprehended as the ratio of the decile of REM to the sum of the deciles of REM and $DACC\_abs$. CIS is client importance, calculated as ratio of total client sales to total sales of the audit firm. All other variables are defined in Appendix C.

Table 33 – Multiple regressions of EM proxies on CINAS for two consecutive years

	Dependent variable: $DACC\_abs_{it}$		Dependent variable: REM_{it}		Dependent variable: $REMvsAEM_{it}$	
	(1) >40 % $CINAS$	(2) >50 % $CINAS$	(3) >40 % $CINAS$	(4) >50 % $CINAS$	(5) >40 % $CINAS$	(6) >50 % $CINAS$
$CINAS_{it}$	0.032	0.030	-0.110	-0.092	-0.036	-0.049
	(1.44)	(1.65)*	(0.19)	(0.17)	(0.64)	(0.93)
DEP_{it}	-0.015	0.163	0.708	4.456	-0.086	-0.430
	(0.44)	(1.43)	(0.57)	(3.96)***	(1.02)	(1.15)
$CINAS_{it} * DEP_{it}$	0.036	-0.233	-1.224	-6.996	0.123	0.669
	(0.61)	(1.46)	(0.51)	(3.18)***	(0.76)	(1.22)
REM_{it}	-0.001	-0.001				
	(0.84)	(0.91)				
$DACC\_abs_{it}$			-0.770	-0.840		
			(0.80)	(0.88)		
$Size_{it-1}$	-0.005	-0.005	-0.045	-0.045	0.012	0.012
	(2.62)***	(2.55)**	(0.57)	(0.57)	(2.33)**	(2.32)**
MB_{it-1}	0.001	0.001	0.023	0.024	-0.005	-0.005
	(0.45)	(0.49)	(0.43)	(0.45)	(1.64)	(1.66)*
$LOSS_{it}$	0.017	0.017	-0.342	-0.342	-0.013	-0.014
	(1.87)*	(1.83)*	(1.92)*	(1.92)*	(0.69)	(0.70)
$Litigation_{it}$	-0.000	-0.000	0.411	0.412	-0.022	-0.022
	(0.08)	(0.05)	(1.67)*	(1.68)*	(1.15)	(1.10)
ACC_{it-1}	-0.023	-0.024	0.028	-0.005	-0.023	-0.019
	(0.61)	(0.65)	(0.08)	(0.02)	(0.51)	(0.43)
CFO_{it}	-0.014	-0.011	6.264	6.313	-0.401	-0.406
	(0.30)	(0.25)	(6.30)***	(6.37)***	(6.01)***	(6.08)***
$Leverage_{it-1}$	-0.016	-0.017	0.271	0.266	0.030	0.031
	(1.08)	(1.10)	(0.63)	(0.61)	(1.07)	(1.09)
ΔTA_{it}	-0.002	-0.000	0.372	0.396	-0.038	-0.040
	(0.12)	(0.02)	(1.90)*	(1.99)**	(2.11)**	(2.37)**
$\Delta EARN_{it}$	0.007	0.006	0.500	0.475	-0.058	-0.054
	(0.48)	(0.42)	(1.18)	(1.11)	(1.44)	(1.27)
ROA_{it}	0.029	0.026	-2.346	-2.401	0.090	0.094
	(0.67)	(0.59)	(2.01)**	(2.12)**	(1.12)	(1.22)
$FirmAge_{it}$	0.002	0.002	-0.019	-0.027	-0.005	-0.004
	(0.61)	(0.54)	(0.10)	(0.15)	(0.32)	(0.27)
$PartChange_{it}$	-0.005	-0.005	0.018	0.027	-0.009	-0.010
	(1.52)	(1.59)	(0.15)	(0.23)	(1.08)	(1.17)
$AudChange_{it}$	0.005	0.006	-0.053	-0.055	0.016	0.017
	(0.63)	(0.65)	(0.24)	(0.25)	(0.78)	(0.80)
$AudOpClean_{it}$	-0.032	-0.031	0.227	0.215	0.008	0.010
	(1.30)	(1.29)	(0.62)	(0.57)	(0.61)	(0.54)
$BoardSize_{it}$	-0.001	-0.001	0.039	0.038	-0.003	-0.003
	(1.38)	(1.40)	(1.05)	(1.00)	(1.04)	(1.03)
$BoardInd_{it}$	-0.034	-0.033	-0.865	-0.848	0.109	0.107
	(1.17)	(1.15)	(0.86)	(0.84)	(1.45)	(1.42)
$VioDCGK_{it}$	0.000	0.000	-0.013	-0.013	0.001	0.001
	(0.65)	(0.72)	(0.70)	(0.69)	(0.58)	(0.61)
$Intercept$	0.175	0.175	0.246	0.269	0.354	0.352
	(3.54)***	(3.54)***	(0.25)	(0.28)	(5.12)***	(5.07)***
$Industry\ FE$	Included	Included	Included	Included	Included	Included
$Year\ FE$	Included	Included	Included	Included	Included	Included
R^2	0.05	0.05	0.16	0.16	0.09	0.09
F	2.63	2.67	5.90	6.10	5.37	5.45
n	1,052	1,052	1,052	1,052	1,052	1,052

*, **, *** Indicate statistical significance at the 0.10, 0.05, and 0.01 levels, respectively, based on two-tailed tests. All test statistics are calculated using OLS with two-way clustering at the firm and time level. We additionally include appropriate fixed effects. Mean VIF for all regressions are considerably below 5. $DACC\_abs$ (REM) is comprehended as the residual from pooled first-stage regressions of Equation 12 (Equation 13, Equation 14 and Equation 17) with two-way clustering. $REMvsAEM$ is comprehended as the ratio of the decile of REM to the sum of the deciles of REM and $DACC\_abs$. $CINAS$ is client importance, calculated as the ratio of client non-audit fees to total fees received from the client. All other variables are defined in Appendix C.

3.7 Additional analyses

3.7.1 Economic bonding, audit service line and the importance of PIEs

We try to present additional evidence and explanation for our findings in chapter 3.6, in particular a reasoning why small and midsized auditors might not be subject to high dependence threats from their publicly listed clients. **Figure 5** contains a matrix showing two importance ratios, e.g., the importance of the audit service line sales for the total sales of the auditor itself ($Ratio\ A$) and the importance of total fees from all PIEs for the total sales of the auditor ($Ratio\ B$).

Figure 5 – Economic bonding and the importance of PIEs

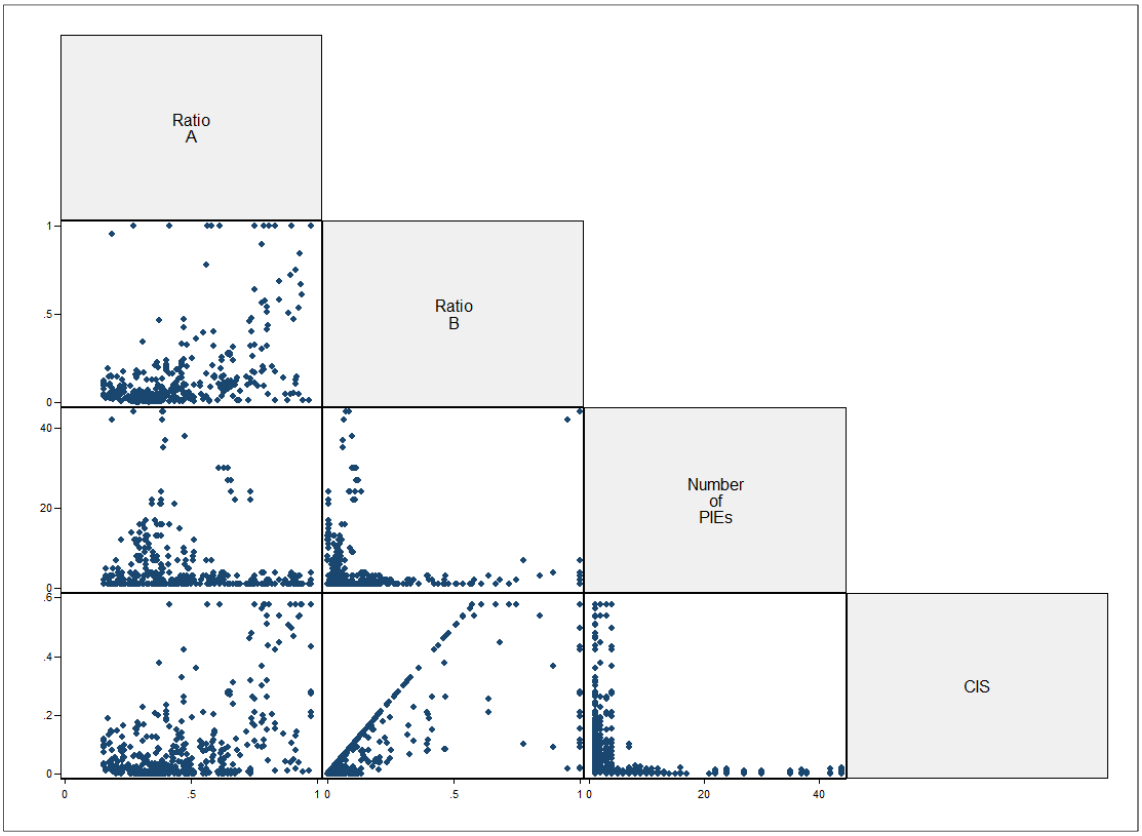


Figure shows a matrix of the distributions of firm-year observations for $Ratio\ A$, $Ratio\ B$, the number of PIEs of the auditor and CIS.

We find that the importance of all PIEs ($Ratio\ B$) is below 50 percent for the largest part of the sample, indicating that the general dependence from fees from PIEs is not at a

critical level.[118] Likewise, while small and midsized audit firms might have only a limited number of PIEs, their overall importance for the total sales ($Ratio\ B$) is not systematically beyond values of 50 percent, which might be considered as a critical value. We also find for $Ratio\ A$ that there are several auditors with low audit service line importance, even though values for CIS are considerable.[119]

Altogether, small and midsized auditors seem to be highly engaged in in the market for non-PIEs, which provides them with a healthy base from there on to operate in the market for PIEs, while they are likewise considerably differentiated along their service lines, e.g., audit, tax & advisory services. To check whether our initial tendency of CIS in the multivariate results of chapter 3.6 is driven by these influencing factors, we add two indicators ($ASIL_{it}$ and $TFIPL_{it}$) for low values of the discussed ratios to our specification for AEM and interact them with CIS.[120] Results are reported in **Table 34**. In particular, we find for both situations (columns (1) & (2)), e.g., low importance of the audit service line and low importance of the PIEs for the total sales of auditor, that our results for CIS are driven by firm-years with auditors, who are located in these parts of the distributions. Hence, both could be a possible explanation for our finding. In column (3), we check for the existence of both effects and only find a significance for the coefficient on $CIS_{it} * TFIPL_{it}$. Based on this finding, we propose that the dominance of reputation and litigation effects is predominantly caused by auditors with high differentiation in markets for PIEs and non-PIEs, giving them a favourable position when it comes to negotiating accounting numbers, thereby increasing audit quality.

[118] Our sample firms account for a mean importance of all PIEs ($Ratio\ B$) of 9.94 percent of the total sales of the audit firm.
[119] However, we note that losing a client's audit fees might entail the loss of non-audit services from the same client, e.g., non-audit fees and audit fees are highly correlated.
[120] $ASIL_{it}$ is a low audit service line importance indicator, calculated as 1 when $Ratio\ A$ is below median $Ratio\ A$, 0 otherwise. $TFIPL_{it}$ is a low total fees from PIEs importance indicator, calculated as 1 when $Ratio\ B$ is below median $Ratio\ B$, 0 otherwise. For reasons of clarity, we exclude the dependence threshold proxies here.

Table 34 – Multiple regressions of AEM on audit service line and PIE client importance

Multiple regressions of alternative AEM on CIS			
Dependent variable: $DACC\_abs_{it}$			
	(1)	(2)	(3)
	$Ratio\ A$	$Ratio\ B$	$Ratio\ A\ \&\ B$
CIS_{it}	-0.019	-0.032	-0.020
	(0.62)	(1.22)	(0.66)
$ASIL_{it}$	0.007		0.006
	(0.87)		(0.91)
$CIS_{it} * ASIL_{it}$	-0.246		-0.218
	(1.82)*		(1.58)
$TFIPL_{it}$		0.007	0.003
		(0.96)	(0.45)
$CIS_{it} * TFIPL_{it}$		-0.870	-0.707
		(3.00)***	(2.30)**
REM_{it}	-0.001	-0.001	-0.001
	(0.99)	(0.87)	(0.94)
$Controls$	Included	Included	Included
$Industry\ FE$	Included	Included	Included
$Year\ FE$	Included	Included	Included
R^2	0.12	0.12	0.13
F	5.76	5.83	5.42
n	1,052	1,052	1,052

*, **, *** Indicate statistical significance at the 0.10, 0.05, and 0.01 levels, respectively, based on two-tailed tests. All test statistics are calculated using OLS with two-way clustering at the firm and time level. We additionally include appropriate fixed effects. Mean VIF for all regressions are considerably below 5. $DACC\_abs$ is comprehended as the residual from pooled first-stage regressions of Equation 12 with two-way clustering. CIS is client importance, calculated as ratio of total client sales to total sales of the audit firm. $ASIL_{it}$ is a low audit service line importance indicator, calculated as 1 when $Ratio\ A$ is below median $Ratio\ A$, 0 otherwise. $TFIPL_{it}$ is a low total fees from PIEs importance indicator, calculated as 1 when $Ratio\ B$ is below median $Ratio\ B$, 0 otherwise. All other variables are defined in Appendix C.

3.7.2 De-facto office and partner level analysis

Auditing research papers have gradually reshaped their level of analysis, since respective audit quality input and output proxies might cause effects on an office or partner level, particularly when studying settings with auditors with large size, e.g., Big 4 settings. Although we focus on small and midsized auditors, size differences are considerable. We try to adapt this approach by re-estimating our main regressions for different levels of analysis in terms of auditor sizes included. We split our sample between auditors with size 1-3 and size 4 following the EU SME definition, resulting in a mean (median) number of clients of 7 (3) for auditors with size 1-3. This level of analysis seems to be comparable to audit office level analyses (e.g., Chung & Kallapur 2003; Sharma et al. 2011). When

we alternatively split the sample between auditors with size 1-2 and size 3-4, mean (median) number of clients of 2 (1) for auditors with size 1-2 is comparable to an audit partner level of analysis. We report our multivariate results for these alternative samples in **Table 35** ($AudSize$ 1-3) and **Table 36** ($AudSize$ 1-2). In **Table 35**, coefficients on our variables of interest show identical signs and partly better significances, indicating that on an office level, our proposed considerations hold, e.g., there seems to be a dominating reputation and litigation effect together with a trade-off from AEM to REM until a certain threshold is met.

What is more, when we take a closer look at a de-facto partner level of analysis in **Table 36**, we find significant results that are once more in line with the already discussed coherences. We conclude that these effects are not caused by any aggregation noise on an audit firm level, but it seems that partners and audit offices pursue increased audit quality for an increasing client importance and do not acquiesce to the client's demands. Their clients tend to switch to a potentially value decreasing trade-off strategy from AEM to REM until a certain critical threshold, where effects tend to reverse.

Table 35 – Multiple regressions of EM proxies on CIS for two consecutive years for AudSize 1-3

	Dependent variable: $DACC\_abs_{it}$		Dependent variable: REM_{it}		Dependent variable: $REMvsAEM_{it}$	
	(1)	(2)	(3)	(4)	(5)	(6)
	>5 % CIS	>10 % CIS	>5 % CIS	>10 % CIS	>5 % CIS	>10 % CIS
CIS_{it}	-0.069	-0.063	-2.054	-2.193	0.192	0.212
	(2.58)**	(1.52)	(1.75)*	(2.14)**	(2.46)**	(2.84)***
DEP_{it}	-0.017	-0.045	-0.323	-0.382	0.070	0.123
	(1.28)	(2.44)**	(0.87)	(0.92)	(2.51)**	(3.03)***
$CIS_{it} * DEP_{it}$	0.107	0.174	2.301	2.675	-0.291	-0.452
	(1.68)*	(1.93)*	(1.30)	(1.39)	(2.42)**	(3.51)***
REM_{it}	0.003	0.003				
	(2.26)**	(2.45)**				
$DACC\_abs_{it}$			1.253	1.210		
			(2.33)**	(2.55)**		
$Size_{it-1}$	-0.012	-0.013	-0.177	-0.184	0.035	0.038
	(3.04)***	(2.98)***	(2.14)**	(2.15)**	(3.92)***	(4.42)***
MB_{it-1}	0.001	0.002	0.090	0.091	-0.013	-0.013
	(1.51)	(1.55)	(1.86)*	(1.85)*	(3.06)***	(3.05)***
$LOSS_{it}$	0.000	-0.001	-0.496	-0.496	0.020	0.022
	(0.00)	(0.07)	(2.32)**	(2.21)**	(0.61)	(0.62)
$Litigation_{it}$	-0.025	-0.027	1.042	1.026	-0.012	-0.007
	(4.47)***	(4.58)***	(3.88)***	(3.86)***	(0.40)	(0.22)
ACC_{it-1}	-0.084	-0.086	-1.371	-1.429	0.111	0.116
	(1.89)*	(1.81)*	(1.91)*	(2.22)**	(2.42)**	(3.53)***
CFO_{it}	-0.154	-0.152	7.897	7.893	-0.461	-0.459
	(0.87)	(0.87)	(9.54)***	(10.03)***	(2.46)**	(2.58)**
$Leverage_{it-1}$	0.020	0.020	0.271	0.298	-0.020	-0.018
	(1.44)	(1.47)	(0.54)	(0.59)	(0.38)	(0.35)
ΔTA_{it}	0.011	0.012	0.267	0.261	-0.070	-0.072
	(0.54)	(0.59)	(0.71)	(0.67)	(1.53)	(1.62)
$\Delta EARN_{it}$	-0.034	-0.035	-0.998	-1.024	0.042	0.047
	(1.23)	(1.23)	(1.58)	(1.73)*	(0.56)	(0.71)
ROA_{it}	0.054	0.053	-3.112	-3.079	0.226	0.218
	(0.32)	(0.32)	(2.73)***	(2.76)***	(1.41)	(1.41)
$FirmAge_{it}$	-0.012	-0.013	0.138	0.130	0.012	0.015
	(1.17)	(1.24)	(0.52)	(0.49)	(0.51)	(0.63)
$PartChange_{it}$	0.004	0.004	-0.074	-0.072	-0.024	-0.026
	(0.69)	(0.84)	(0.30)	(0.29)	(1.32)	(1.34)
$AudChange_{it}$	-0.023	-0.025	-0.126	-0.101	0.043	0.042
	(1.34)	(1.36)	(0.42)	(0.33)	(1.33)	(1.28)
$AudOpClean_{it}$	-0.019	-0.021	0.349	0.354	0.046	0.048
	(0.69)	(0.83)	(0.84)	(0.80)	(1.11)	(1.02)
$BoardSize_{it}$	-0.003	-0.003	0.018	0.017	0.003	0.003
	(1.36)	(1.32)	(0.34)	(0.35)	(0.56)	(0.56)
$BoardInd_{it}$	-0.044	-0.039	-0.299	-0.261	0.024	0.011
	(1.24)	(1.10)	(0.28)	(0.25)	(0.23)	(0.11)
$VioDCGK_{it}$	-0.001	-0.001	-0.011	-0.009	0.003	0.003
	(0.60)	(0.60)	(0.52)	(0.41)	(1.13)	(1.00)
$Intercept$	0.305	0.320	0.553	0.576	0.067	0.029
	(4.39)***	(4.21)***	(0.58)	(0.52)	(0.59)	(0.23)
$Industry\ FE$	Included	Included	Included	Included	Included	Included
$Year\ FE$	Included	Included	Included	Included	Included	Included
R^2	0.14	0.14	0.32	0.32	0.16	0.16
F	3.79	3.67	9.84	10.54	4.65	4.67
n	499	499	499	499	499	499

*, **, *** Indicate statistical significance at the 0.10, 0.05, and 0.01 levels, respectively, based on two-tailed tests. All test statistics are calculated using OLS with two-way clustering at the firm and time level. We additionally include appropriate fixed effects. Mean VIF for all regressions are considerably below 5. $DACC\_abs$ (REM) is comprehended as the residual from pooled first-stage regressions of Equation 12 (Equation 13, Equation 14 and Equation 17) with two-way clustering. $REMvsAEM$ is comprehended as the ratio of the decile of REM to the sum of the deciles of REM and $DACC\_abs$. CIS is client importance, calculated as ratio of total client sales to total sales of the audit firm. $AudSize$ is a categorical variable based on the EU SME definition from May 2003. All variables are defined in Appendix C.

Table 36 – Multiple regressions of EM proxies for two consecutive years for AudSize 1-2

	Dependent variable: $DACC\_abs_{it}$		Dependent variable: REM_{it}		Dependent variable: $REMvsAEM_{it}$	
	(1)	(2)	(3)	(4)	(5)	(6)
	>5 % CIS	>10 % CIS	>5 % CIS	>10 % CIS	>5 % CIS	>10 % CIS
CIS_{it}	-0.124	-0.128	-1.910	-1.699	0.189	0.153
	(2.62)***	(1.89)*	(1.88)*	(2.57)**	(2.72)***	(2.59)**
DEP_{it}	-0.019	-0.049	-0.412	-0.410	0.060	0.104
	(1.07)	(2.05)**	(1.11)	(1.12)	(1.55)	(1.97)**
$CIS_{it} * DEP_{it}$	0.150	0.229	2.125	2.083	-0.274	-0.362
	(1.95)*	(2.09)**	(1.26)	(1.33)	(2.67)***	(2.23)**
REM_{it}	0.002	0.002				
	(0.98)	(0.98)				
$DACC\_abs_{it}$			1.034	0.963		
			(1.07)	(1.07)		
$Size_{it-1}$	-0.009	-0.011	-0.254	-0.282	0.031	0.038
	(2.04)**	(1.91)*	(1.92)*	(2.19)**	(1.92)*	(2.50)**
MB_{it-1}	0.001	0.001	0.053	0.056	-0.008	-0.009
	(0.40)	(0.47)	(0.85)	(0.87)	(1.29)	(1.31)
$LOSS_{it}$	0.010	0.008	-0.019	-0.022	-0.034	-0.031
	(0.84)	(0.65)	(0.09)	(0.10)	(0.80)	(0.66)
$Litigation_{it}$	-0.023	-0.026	1.316	1.303	-0.081	-0.076
	(2.27)**	(2.51)**	(4.06)***	(3.85)***	(2.03)**	(1.81)*
ACC_{it-1}	-0.113	-0.117	-0.440	-0.536	0.150	0.158
	(1.76)*	(1.79)*	(0.42)	(0.54)	(1.34)	(1.71)*
CFO_{it}	-0.186	-0.179	7.438	7.387	-0.387	-0.385
	(1.54)	(1.55)	(6.79)***	(7.33)***	(2.11)**	(2.23)**
$Leverage_{it-1}$	0.040	0.043	-0.257	-0.257	0.019	0.022
	(1.75)*	(2.01)**	(0.38)	(0.38)	(0.21)	(0.25)
ΔTA_{it}	0.038	0.040	0.479	0.496	-0.103	-0.109
	(1.15)	(1.11)	(0.75)	(0.78)	(1.99)**	(2.23)**
$\Delta EARN_{it}$	-0.110	-0.111	-0.218	-0.296	0.088	0.096
	(1.88)*	(1.80)*	(0.17)	(0.23)	(0.99)	(1.15)
ROA_{it}	0.098	0.091	-2.421	-2.278	0.030	0.019
	(0.75)	(0.74)	(1.34)	(1.30)	(0.17)	(0.11)
$FirmAge_{it}$	-0.020	-0.022	0.039	0.026	0.035	0.040
	(1.27)	(1.40)	(0.14)	(0.09)	(1.11)	(1.20)
$PartChange_{it}$	0.005	0.006	-0.032	-0.022	-0.024	-0.027
	(0.46)	(0.59)	(0.18)	(0.12)	(2.32)**	(2.72)***
$AudChange_{it}$	-0.001	-0.004	-0.309	-0.277	0.016	0.020
	(0.05)	(0.14)	(1.61)	(1.32)	(0.36)	(0.44)
$AudOpClean_{it}$	-0.029	-0.031	0.275	0.244	0.053	0.061
	(0.80)	(0.94)	(0.59)	(0.49)	(0.98)	(1.07)
$BoardSize_{it}$	-0.002	-0.002	0.103	0.100	0.002	0.002
	(0.71)	(0.55)	(1.05)	(1.03)	(0.20)	(0.20)
$BoardInd_{it}$	0.006	0.015	-0.680	-0.640	-0.067	-0.081
	(0.17)	(0.44)	(0.50)	(0.47)	(0.47)	(0.55)
$VioDCGK_{it}$	-0.001	-0.001	0.002	0.005	0.002	0.001
	(0.28)	(0.28)	(0.10)	(0.23)	(0.47)	(0.41)
$Intercept$	0.251	0.273	1.590	1.834	0.125	0.053
	(3.59)***	(3.19)***	(0.94)	(1.01)	(0.65)	(0.28)
$Industry\ FE$	Included	Included	Included	Included	Included	Included
$Year\ FE$	Included	Included	Included	Included	Included	Included
R^2	0.16	0.18	0.35	0.35	0.18	0.19
F	2.50	2.39	5.30	5.56	3.55	3.60
n	287	287	287	287	287	287

*, **, *** Indicate statistical significance at the 0.10, 0.05, and 0.01 levels, respectively, based on two-tailed tests. All test statistics are calculated using OLS with two-way clustering at the firm and time level. We additionally include appropriate fixed effects. Mean VIF for all regressions are considerably below 5. $DACC\_abs$ (REM) is comprehended as the residual from pooled first-stage regressions of Equation 12 (Equation 13, Equation 14 and Equation 17) with two-way clustering. $REMvsAEM$ is comprehended as the ratio of the decile of REM to the sum of the deciles of REM and $DACC\_abs$. CIS is client importance, calculated as ratio of total client sales to total sales of the audit firm. $AudSize$ is a categorical variable based on the EU SME definition from May 2003. All variables are defined in Appendix C.

3.7.3 Additional analysis of extreme EM

To address the claim that EM is not a systematic phenomenon, we follow Greiner et al. (2017) and use only top quintile EM observations as a proxy where EM is used to foster opportunistic goals. We therefore introduce two dummy variables ($Q1DA$, $Q1REM$) as dependent variables of **Equation 24-Equation 25**. These are equal to 1 if the EM proxy ($DACC\_abs_{it}$, REM_{it}) is located in the top quintile of the industry-year distribution, and 0 otherwise. Results from this alteration are reported in **Table 37**.[121] We find that all of our major results remain the same. Still, we consistently find that small and midsized audit firms might have higher reputation and litigation risk incentives, which increases audit quality, while these are superimposed by economic bonding effects when the importance of the client reaches the hypothesized thresholds, particularly when the bond maintains for a certain period of time, e.g., two or three years. We also test for the robustness of no significances of $CINAS$ when using extreme EM proxies. Again, there are no significant coefficients for our client importance proxies based on non-audit service importance.

[121] We use logit estimations with two-way clustering and respective fixed effects to remain consistent and address our change in specification. Greiner et al. (2017) use signed discretionary accruals for their analysis, because Abbott et al. (2006) highlight the asymmetry of the fee response to income increasing or decreasing EM. For our purposes, we regard both income increasing and decreasing EM as a distortion of financial reporting quality, which is why we use distributions of absolute values of discretionary accruals.

Table 37 – Multiple regressions of extreme AEM and REM on CIS

	>5 % *CIS*			**>10 % *CIS***			**>15 % *CIS***		
	Dependent variable: Dummy for the extreme quintile of the distribution of the positively scaled absolute value of discretionary accruals: $Q1DACC\_abs_{it}$								
	(1) one year	(2) two years	(3) three years	(4) one year	(5) two years	(6) three years	(7) one year	(8) two years	(9) three years
CIS_{it}	-6.659	-5.568	-5.839	1.222	-5.063	-5.557	-2.985	-3.683	-3.793
	(0.80)	(2.11)**	(2.25)**	(0.59)	(4.75)***	(5.35)***	(1.19)	(2.01)**	(2.53)**
DEP_{it}	-0.033	-0.236	-0.261	-0.604	-0.383	-0.267	-0.187	-0.449	-0.461
	(0.11)	(0.82)	(0.73)	(0.95)	(0.60)	(0.35)	(0.20)	(0.46)	(0.40)
$CIS_{it} * DEP_{it}$	5.670	5.511	6.008	-0.659	5.569	5.997	2.530	4.338	4.680
	(0.74)	(2.23)**	(2.75)***	(0.23)	(2.40)**	(2.91)***	(0.73)	(1.36)	(1.32)
$Q1REM_{it}$	0.231	0.249	0.249	0.246	0.251	0.244	0.241	0.247	0.243
	(0.87)	(0.92)	(0.92)	(0.92)	(0.91)	(0.89)	(0.90)	(0.90)	(0.88)
Controls	Included	Included	Included	Included	Included	Included	Included	Included	Included
Industry FE	Included	Included	Included	Included	Included	Included	Included	Included	Included
Year FE	Included	Included	Included	Included	Included	Included	Included	Included	Included
R^2	0.03	0.04	0.04	0.04	0.04	0.04	0.03	0.04	0.04
chi^2	34.71	37.67	39.53	35.50	39.59	41.72	35.66	37.00	38.68
n	1,052	1,052	1,052	1,052	1,052	1,052	1,052	1,052	1,052
	Dependent variable: Dummy for the extreme quintile of the distribution of the comprehensive positively scaled standardized *REM* proxy: $Q1REM_{it}$								
CIS_{it}	-4.059	1.491	0.859	3.022	1.326	0.043	3.348	1.243	0.315
	(0.29)	(1.19)	(0.59)	(0.74)	(1.10)	(0.02)	(1.00)	(1.04)	(0.16)
DEP_{it}	0.616	0.803	0.772	0.856	1.070	0.789	0.764	1.129	0.445
	(1.83)*	(2.42)**	(2.13)**	(1.69)*	(1.88)*	(1.48)	(1.55)	(1.58)	(0.54)
$CIS_{it} * DEP_{it}$	2.113	-4.497	-3.678	-5.505	-4.988	-2.427	-5.550	-4.831	-1.770
	(0.15)	(2.60)***	(1.85)*	(1.19)	(2.09)**	(0.82)	(1.41)	(1.69)*	(0.56)
$Q1DACC\_abs_{it}$	0.151	0.169	0.169	0.167	0.173	0.164	0.162	0.173	0.167
	(0.64)	(0.69)	(0.69)	(0.69)	(0.70)	(0.67)	(0.68)	(0.70)	(0.67)
Controls	Included	Included	Included	Included	Included	Included	Included	Included	Included
Industry FE	Included	Included	Included	Included	Included	Included	Included	Included	Included
Year FE	Included	Included	Included	Included	Included	Included	Included	Included	Included
R^2	0.05	0.05	0.05	0.05	0.05	0.05	0.05	0.05	0.04
chi^2	45.41	47.65	46.02	45.63	46.88	46.91	44.37	45.45	40.29
n	1,052	1,052	1,052	1,052	1,052	1,052	1,052	1,052	1,052

*, **, *** Indicate statistical significance at the 0.10, 0.05, and 0.01 levels, respectively, based on two-tailed tests. All test statistics are calculated using logit with two-way clustering at the firm and time level. We keep our fixed effects and controls as stated in section 3.4.2. Mean VIF for all regressions are considerably below 5. $Q1DACC\_abs$ ($Q1REM$) is a dummy for observations in the extreme quintile of the respective distribution of $DACC\_abs$ (REM), which is the residual from cross-sectional industry-year specific first-stage regressions of Equation 12 (Equation 13, Equation 14 and Equation 17). We use inverted REM for the quintile calculation. CIS is client importance, calculated as ratio of total client sales to total sales of the audit firm. All other variables are defined in Appendix C.

3.7.4 Audit committee moderating effects

Sharma et al. (2011) provide insights into the possible moderating role of an audit committee meeting best practice requirements. To address this issue, we further control for the existence of such an audit committee as well as possible moderation of our effects. DCGK paragraph 5.3.2 requires the establishment of an audit committee, which has the duty to appoint and monitor the auditor, in particular with regard to dependence threats. Furthermore, DCGK paragraph 5.3.2 lays down requirements for the chair of the audit committee, e.g., (1) being an independent member, (2) being an accounting expert, (3) not having been an executive member of the board recently (last three years) nor (4) being the chair of the supervisory board.[122] However, as mentioned above, the DCGK is a voluntary commitment to high-quality corporate governance, while Art. 161 AktG requires an annual declaration by the supervisory and executive board with regard to the accordance of the corporate governance of the firm with the DCGK. We follow Sharma et al. (2011) and add a dummy variable (AC_{it}) for the existence and best practice conformity of the audit committee to our specification and include interaction with our CI variables to study the moderating effects of the audit committee. Results are reported in **Table 38**.

We find that our results are very robust to the inclusion of a possible moderation by a best practice audit committee. Even more, the significance of our results improves. In particular, this alteration provides additional interesting insights, e.g., in columns (3)-(4) we find that the existence of a best practice audit committee seems to improve the financial reporting quality in terms of a decrease in the use of REM. This is surprising taking into account the original task of the audit committee to mediate in accounting decisions. What is important, audit committee's duties have to be carried out by the board if there is no audit committee. Based on a mean $BoardInd_{it}$ of 65.16 percent and a mean $BoardInd_{it}$ of 63.24 percent when there is no audit committee, best practice audit committees might monitor beyond accounting based EM and even restrict use of REM, while the highly independent boards might do a comparably good job in monitoring AEM (e.g., Klein 2002; Bradbury et al. 2006). This finding is supported by a look at the results in columns

[122] These requirements are highly aligned with the ones used by Sharma et al. (2011), who find that a best practice audit committee can help secure independence.

(5)-(6) for $REMvsAEM_{it}$, while the trade-off relation for our hypothesized thresholds remains.

Table 38 – Multiple regressions of alternative EM proxies on CIS and AC for two consecutive years

	Dependent variable: $DACC\_abs_{it}$		Dependent variable: REM_{it}		Dependent variable: $REMvsAEM_{it}$	
	(1) >5 %CIS	(2) >10 % CIS	(3) >5 % CIS	(4) >10 % CIS	(5) >5 % CIS	(6) >10 % CIS
CIS_{it}	-0.116	-0.103	-1.326	-1.616	0.171	0.142
	(32.03)***	(2.79)***	(1.23)	(2.17)**	(2.33)**	(1.65)
AC_{it}	0.006	0.006	0.435	0.418	-0.047	-0.049
	(0.86)	(0.84)	(2.04)**	(1.98)**	(2.78)***	(2.80)***
$CIS_{it} * AC_{it}$	-0.011	-0.023	-2.014	-0.319	0.100	0.293
	(0.23)	(0.32)	(1.02)	(0.14)	(0.57)	(0.99)
DEP_{it}	-0.019	-0.044	-0.232	-0.353	0.055	0.098
	(2.14)**	(2.79)***	(0.71)	(0.87)	(1.95)*	(2.40)**
$CIS_{it} * DEP_{it}$	0.155	0.210	2.147	2.914	-0.299	-0.377
	(3.11)***	(3.03)***	(1.29)	(1.80)*	(2.21)**	(2.95)***
$CIS_{it} * DEP_{it} * AC_{it}$	-0.030	-0.025	-0.241	-2.352	0.097	-0.108
	(0.45)	(0.34)	(0.11)	(1.09)	(0.47)	(0.36)
REM_{it}	-0.002	-0.002				
	(1.09)	(1.17)				
$DACC\_abs_{it}$			-0.495	-0.513		
			(1.07)	(1.14)		
$Controls$	Included	Included	Included	Included	Included	Included
$Industry\ FE$	Included	Included	Included	Included	Included	Included
$Year\ FE$	Included	Included	Included	Included	Included	Included
R^2	0.12	0.13	0.31	0.31	0.15	0.15
F	5.16	5.18	16.54	17.21	8.16	8.09
n	1,052	1,052	1,052	1,052	1,052	1,052

*, **, *** Indicate statistical significance at the 0.10, 0.05, and 0.01 levels, respectively, based on two-tailed tests. All test statistics are calculated using OLS with two-way clustering at the firm and time level. We additionally include appropriate fixed effects. Mean VIF for all regressions are considerably below 5. $DACC\_abs$ (REM) is comprehended as the residual from pooled first-stage regressions of Equation 12 (Equation 13, Equation 14 and Equation 17) with two-way clustering. $REMvsAEM$ is comprehended as the ratio of the decile of REM to the sum of the deciles of REM and $DACC\_abs$. CIS is client importance, calculated as ratio of total client sales to total sales of the audit firm. AC_{it} is existence of a best practice audit committee, calculated as 1 when firms have an audit committee in accordance with the DCGK 5.3.2, 0 otherwise. All other variables are defined in Appendix C.

3.7.5 Robustness analyses

To address the problem of biased EM proxies, we use the alternative models discussed in section 2.2. For AEM, we apply the adapted performance-matched Jones model and the McNichols (2002) model, while for REM, we apply the Gunny (2010) models. We additionally use going concern opinions as a dependent variable in our regressions to be able to study another frequently used audit quality output based measure. All results from these alterations are qualitatively the same and remain untabulated.

Another common criticism is focusing on residual proxies with insufficient separation of non-discretionary and discretionary accrual parts. This could lead to false inferences from

second-step regressions when omitted correlated non-discretionary variables exist, which explain considerable variation in the discretionary proxies because these still include non-discretionary variation. We therefore use one-step models with non-discretionary and discretionary controls and our variables of interest and estimate these regressions again using pooled OLS with two-way clustering on the firm and year level. Decisively, our results are very robust to this alteration and further convince us of the validity of our analysis.[123] We also try to address the possible simultaneity of the decisions to pursue AEM and REM. Although REM has to be performed throughout the fiscal year, managers might try to include expected AEM possibilities in their decision to pursue REM. Therefore, we set up a simultaneous equation model and apply 3SLS to comprehend residuals from the simultaneous equations. These changes in residuals leave us with the same results as already presented, which is why we assume the problem of simultaneity in our setting to be considerably low.[124]

Likewise, our two-way clustering on the year level on the second stage might suffer from low number of clusters (Cameron et al. 2011). Therefore, we check for the robustness of our results when we cluster only on one level (firm) or on the two levels of firm and audit firm. We keep our fixed effects setting based on our discussion above. Results from these untabulated tests remain qualitatively the same and quantitatively almost equal, affirming our findings in chapter 3.6 and 3.7.

What is more, we further address the non-linearity of our relationship and alternatively estimate a regression using a linear and squared regressor for our client importance proxy CIS. In addition, we follow Chi et al. (2011) and alternatively test these specifications with the natural logarithm of the total fees from a client and the total non-audit fees from a client as proxies for auditor independence (e.g., DeFond et al. 2000; Francis 2004). We find significant coefficients for these EM proxies, implying that the opposing incentives superimpose each other diversely in different parts of the distributions. Signs of the coefficients affirm our results presented in chapter 3.6 and 3.7.

[123] Again, we do not report the respective tables for these robustness regressions.
[124] Since our results remain qualitatively the same, we do not tabulate the regressions.

To test the robustness of our results with respect to type I errors, we extend the specification and add confounding compensation controls for the executive and supervisory part of the board (compensation magnitude per board members, degree of short- and long-term variable compensation of executives). We further address possible type I errors caused not by the omission but the inclusion of collider or mediator variables in our empirical model. We therefore exclude $PartChange_{it}$, $AudChange_{it}$ and $AudOpClean_{it}$ since these variables might collide with our variables of interest and dependent variables. However, untabulated regressions show that this alteration does not change our results, nor does stepwise inclusion.

3.8 Conclusion

This study analyses the effect of client importance on the magnitude and type of EM used in PIE client firms of the German small and midsized auditor market. In particular, we try to disentangle the total effect by introducing critical thresholds for client importance, thereby separating the effect for clients with pronounced economic bonding. We apply two precise measures for client importance, e.g., client importance from client's total fees for audit and non-audit services and client importance from client's non-audit services, while we use the total sales of the auditor from all clients and services in the denominator. Furthermore, we study a possible trade-off relationship between AEM and REM, check for the moderation effects of a best practice audit committee, check for possible changes in the results in de-facto office and partner level analyses and provide a reasoning for our results by analyzing the importance of PIEs for the small and midsized auditors in more detail.

For normal levels of client importance from total fees, our results indicate that increasing client importance decreases AEM and increases REM. However, when client importance reaches the 5%, 10% or 15% threshold for one, two or three consecutive years, the results show that effects reverse, particularly for increasing time-period and threshold. Hence, the opposing incentives related to client importance dominate in different parts of the distribution. Below the hypothesized thresholds, higher dependence effects tend to be overcompensated by higher reputation and litigation effects, which results in an improved

audit quality that forces clients to trade-off AEM with REM. When the hypothesized thresholds are met, the magnitude of the diverse incentives reverse and hence impairs audit quality, leaving more room for AEM and a trade-off back from REM to AEM. For client importance measured by the percentage of non-audit fees to total client's fees (non-audit fees ratio), we find no significant effects on AEM or REM. We conclude that both proxies measure diverse aspects of the relationship between client and the auditor, while the client importance from total client fees is clearly the best indicator for client importance. Our study furthermore shows that small and midsized auditors with only a few number of PIE clients do not necessarily deliver a low audit quality, but may have a solid foundation of non-PIE clients, from there they can deliver high audit quality without being too dependent on the respective publicly listed client. We also show that our results are not dependent on omitted correlated variables like a best practice audit committee, while including this aspect sheds some light on to how shareholders can effectively limit the use of REM and AEM. Furthermore, our results hold in both de-facto office and de-facto partner level analyses and do not seem to be subject to design choices or caused by biased EM proxies.

Our study contributes to the existing research in the following ways. First, existing studies about the influence of client importance on EM often noisily measure client importance, particularly when drawing inferences about the non-Big N auditors, which might not have a considerable number of non-PIE clients. We complement these studies by using more precise proxies for client importance for a setting of small and midsized auditors. We show that this alteration makes an important difference when studying the economic bonding for this auditor market segment and therefore provide a fresh look at small and midsized auditors when it comes to their eligibility for PIE client engagements. Second, we investigate regulatory-driven dependence thresholds to study the non-linearity of the effects caused by opposing incentives associated with client importance. This might also be useful for regulatory bodies to assess the validity of the respective rules and standards. Third, we consider the trade-off between the two different types of EM, raising the awareness of the option for clients to circumvent high quality audits by simply using more REM, which is harder to detect and/or challenge for any auditor during the audit process.

Here, we provide another interesting result showing the contribution of a best practice audit committee.

Naturally, our analysis has some limitations. First, studies on EM decisively hinge on the question of whether the proxies for EM have been collocated using the right models and estimation procedures to ensure the discretionary components are not biased. We try to use the most sophisticated approaches to minimize biases, although there is no guarantee for that. What is more, we control for the commonly used confounding controls to explain the variation of the discretionary EM proxies. Nevertheless, although comparable to the level of prior studies, the explanatory power of our multivariate regressions is only moderate, which is why considerable influence of omitted variables might be possible.[125] We also have a considerable amount of excluded observations, e.g., a sample selection bias cannot be cancelled out. In particular, financially distressed firms do not publish their financial statements during insolvency proceedings, while these observations might be indeed of above-average interest, since they might be more likely to use EM.

[125] However, a high heterogeneity of our sample firms could be another explanation for a moderate r-squared.

4 Differentiated analysis of the influence of corporate social responsibility on earnings management and the role of corporate governance[126]

The need for uniformity in the multitude of sustainability standards. In his speech, the IASB chair Hoogervorst also comments on the non-financial information about social and environmental issues:

> "That is not our area of expertise. […] Our remit is, and will remain, financial reporting – with focus on the participants in the capital markets. That is investors and potential creditors." (Hoogervorst 2017)

Following the outline by Penman and Sougiannis (1998), our aggregated construct X included all attributes of a firm that capture value-creating activities. Hoogervorst (2017) deliberately acknowledges the fact that shareholders, and not only stakeholders, are interested in sustainability issues as means of creating value in certain, if not all industries. Likewise, since Corporate Social Responsibility has developed more into a vehicle of systematic non-financial opportunities and risks to foster reputation and create value (e.g., Gastinger & Gaggl 2015), there is need for the analysis of the interrelation between CSR, sustainability reporting and financial reporting, since all can be means of considerable discretion and reputation enhancement strategies (e.g., Jasch 2015). What is more, Graham et al. (2005) implicitly show in their study when asking about meeting certain earnings benchmarks that CEOs/CFOs use discretion to foster earnings numbers to build external reputation and credibility with the market (see **Figure 6**) as two of the key incentives.

[126] This chapter is based on a 2017 version of a working paper titled "Corporate social responsibility and earnings management – a differentiated view on CSR incentives and the role of corporate governance". The reasoning, results, and interpretations of this might change after the submission of this thesis and the completion of the doctoral degree. The most recent version of this study is available upon request. Please do not cite this working paper without permission.

Figure 6 – Importance of certain incentives behind EM from survey by Graham et al. (2005)

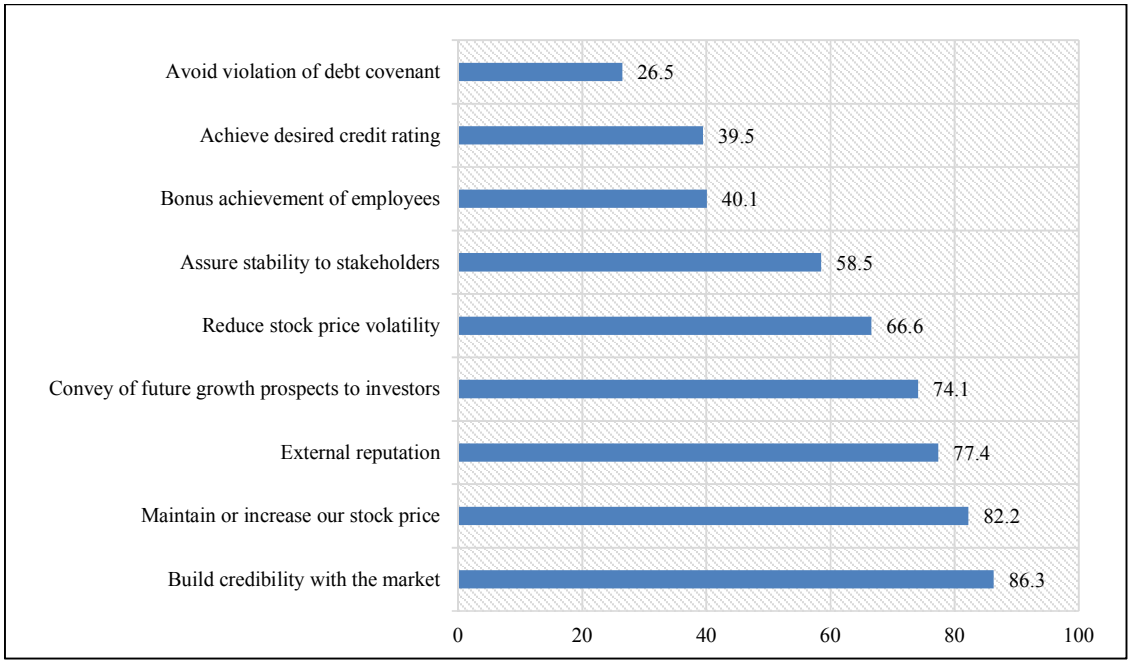


Figure shows results for 304-307 survey responses to the question: "Meeting earnings benchmarks helps…" based on a survey of 401 financial executives. Numbers show percentages of respondents agreeing or strongly agreeing with the respective statement. Source of information: Graham et al. (2005), p. 26-27.

Chapter 4 tries to present an extensive analysis of the coherence between CSR performance and the quality of earnings to disentangle the existent incentives and effects, particularly with respect to reputation incentives, while integrating the influence of Corporate Governance (CG) as the environmental characteristic shaping the quality of financial information.

4.1 Introduction

The importance of disclosing adequate financial information has led to major changes in the nature and scope of reporting. Concerning this adequacy, the interests and needs of all stakeholders play an important role since they consider the presented information in their decisions. As a result, the requirements on corporate disclosure have diversely evolved over time. The endeavours for CSR and well-functioning CG have been two important influencing factors since they are inevitable characteristics and fields of commitment for every entity. First, society expects firms to be ethical and socially responsible to a certain degree, which similarly influences and extends the disclosure of companies. Analysing the influence in more detail, the actual motivations for pursuing CSR can be diverse. Some firms may want to masquerade themselves as socially responsible when they actually are not to improve their image, whereas others may act out of intrinsic motivation. The difference in incentives also shapes the relationship between CSR engagement and financial reporting, particularly with regard to earnings management (EM) and deciding about the various types of EM. Reputational and intrinsic CSR incentives may lead to less accrual-based EM (AEM) and therefore accounting controversies, though the former principally tries to prevent them from affecting reputation. This may have an influence on the use of real activities EM (REM). Second, CG has been highlighted as a firm's moral legitimacy in relation to stakeholders through major regulations, e.g., the Sarbanes-Oxley Act of 2002 (SOX) for the USA or the CG framework green paper of 2011 for the EU. Although there are many influencing factors in the governance of a company, corporate board functions and structures play an integral role (e.g., Scherer & Palazzo 2011; Zhang et al. 2013). Effective CG to ensure proper financial reporting thus means the efficient collocation and functioning of boards to limit EM. Firms must decide about how many internal and external members ensure proper monitoring as well as what personnel characteristics are important to strengthen the viability of disclosure and emphasize the interests of all stakeholders (e.g., Johnson & Greening 1999; Webb 2004; Dalton & Dalton 2010). Although this is straightforward for AEM, business alterations for REM reasons are likely to be more difficult to detect and prohibit.

Altogether, there is a concurrent influence of CSR and CG on accounting information, which either increases or decreases both types of EM in favour or at the expense of the various stakeholders. The overall direction depends on the relevant incentives behind the EM strategies. Prior research on CSR claims that CSR incentives are important (e.g., Kim et al. 2012; Bozzolan et al. 2015), which is why we try to disentangle these in more detail and regard them as a possible moderator of the success of CG in limiting EM. In particular, we include the trade-off decision between AEM and REM. Intrinsic CSR may highlight moral integrity and therefore foster the functioning of CG, whereas reputational CSR may entail a more opportunistic use of EM. Good CG can prevent AEM, whereas there may be a shift in EM activities towards REM, thus potentially threatening the long-term objectives of stakeholders.

To determine the various relationships, this study examines whether firms that act socially responsible use lower or higher EM, though analysing the related CSR incentives in a differentiated manner. Further, we take a closer look at whether CSR incentives moderate the influence of CG characteristics on EM. In particular, we focus not only on the respective magnitude but also on the trade-off relationship between the two types of EM. Therefore, this study uses an ASSET4EU ESG data set, which contains accounting as well as CSR and CG data of the most important European companies. For the sample period of 2005-2014, the findings show that firms with a higher performance in CSR engage in different levels of EM and that incentives play an important role. We show that more intrinsically motivated firms use lower total EM and trade off shareholder hostile REM with AEM. We further elaborate that firms with somewhat more reputational CSR incentives use lower AEM, though minimizing reputational damage through accounting controversies. However, they use considerably more REM and trade off AEM with REM, thereby minimizing controversies while maintaining high earnings that further foster their image. Additionally, we provide interesting results for the effect of board independence on earnings quality in terms of lower AEM and/or REM, taking into account the respective CSR incentives.

This study contributes to the actual strand of literature in the following ways. First, the investigation illuminates the question of whether CSR is positively or negatively linked to EM. To date, the answers to this question have not been satisfying. Many studies find mixed results, which is why further research is essential. In addition, we try to develop previous approaches by further disentangling the incentives behind CSR. Thus, we provide additional insights about what drives firms and their decision in the respective situation. Second, the study provides insights concerning the relationship between CG characteristics and EM. Here, again, the previous findings are mixed and require further research. Third, we connect both research areas and elaborate significant moderating effects of CSR incentives on the relationship between CG characteristics and EM, which can be of major importance for both the CSR and CG areas of research. Our results can also be of interest for both standard setters and regulators interested in the earnings quality of firms with various types of CSR and CG. It also presents various incentives and control variable links that should be useful. Fourth and finally, stakeholders can use the results to obtain credibility for possible future interaction.

The remainder of this part of the doctoral thesis is structured as follows. The next section addresses the theoretical background, related literature and hypothesis development. We introduce our research design in section 4.3, and section 4.4 presents the results of the study. Section 4.5 conducts additional analyses and robustness tests; section 4.6 concludes, provides some remarks and highlights limitations.

4.2 Background, related literature and research hypotheses

4.2.1 Corporate social responsibility and earnings management – a differentiated view

Accounting research has studied EM in a wide variety of ways. Entities use various types of EM in the process of financial reporting and significantly affect the presented picture. However, this picture is highly relevant for contractual outcomes and decisions that rely on the disclosed information. The asymmetry between inside directors and outside stakeholders

plays an important role when studying fiscal reports because it makes selfish behaviour a considerable problem. Therefore, financial statements may reflect the underlying conditions inaccurately when EM is conducted (e.g., Leuz et al. 2003). Despite these limiting factors for the quality of financial information, the accounting literature reports extensive evidence of the use of managed numbers in stakeholders' actions, e.g., successful EM.

More recently, accounting studies have taken great interest in CSR. Society has gradually become more sophisticated about how entities should interact with all types of stakeholders, e.g., investors, customers, suppliers and the respective governments. Regarding this sophistication, presently, firms are expected to act and communicate in a manner that can be regarded as ethical and commensurate with the various interests of all stakeholders (e.g., Reynolds & Yuthas 2008; Shafer 2015). With several high-profile theoretical and analytical research contributions about CSR and its qualities being available (e.g., Carroll 1979; Jones 1995; McWilliams & Siegel 2001; Garriga & Melé 2004; Porter & Kramer 2006; Mackey et al. 2007), Carroll (1979) prompts enterprises to dispute with aspects apart from profitability and maximizing shareholder value. He regards legal, ethical and philanthropic responsibility to be important amplifications for a firm's overall success. These comprise actions that foster social goods that go beyond a firm's interest and legal requirements (e.g., McWilliams & Siegel 2001). Jones (1995) notes that corporations can signal credible interest in trustworthy and cooperative interaction with all stakeholders through CSR activities. However, a holistic approach to CSR will also influence accounting decisions (e.g., Victor & Cullen 1988; Atkins 2006; Thornton 2008; Hong & Andersen 2011). Intrinsic motivation strives for transparency and reliability as fundamental parameters in reporting. Since transparency can be viewed as a function of whether EM is used, firms that characterize themselves as socially responsible should engage less due to their intrinsic motivation.[127] Consequently, there should be a de-

[127] In line with this consideration, the common admission is that EM reduces the information content of the reporting and quality of earnings. This assumption is disputable, particularly in cases of incorrect measurement. Moreover, the inadequacy of accounting systems limits the ways in which firms can state a true and fair earnings picture. They can also signal this picture by using EM. However, inadequacy most likely occurs on an industry-specific basis. We mainly rule out this variation by incorporating respective industry fixed-effects in our analyses. Altogether, the study follows Dechow et al. (2010) and considers the EM variation apart from signalling as quality reducing.

crease in EM with increasing amounts of CSR engagement. The related literature has predominantly studied this negative relationship, often called the "myopia avoidance hypothesis".

Chih et al. (2008) find results that support a negative coefficient of CSR on income smoothing and loss avoidance in their international setting, whereas the overall outcomes are mixed. Calegari et al. (2010) state a negative coherence between CSR and AEM and, therefore, enhanced reporting quality in their paper using USA firms. Hong and Andersen (2011) and Kim et al. (2012; KPW) extend these considerations and take REM into account, whereas KPW and, in particular, Bozzolan et al. (2015; BFMM) focus on a possible trade-off relationship between the different types of EM and their coherence with CSR. KPW find extensive proof of a lower use of AEM and REM when CSR performance is high as well as a general trade-off relationship between both types of EM. They also find that CSR firms are less likely to be the subject of SEC investigations on GAAP violations in Accounting and Auditing Enforcement Releases (AAER). Their results hold for substitution of the KLD measure by Domini 400 Social Index affiliation and the separate consideration of the strengths and weaknesses that are part of the KLD measure. BFMM analyse the trade-off relationship between AEM and REM in more detail, taking into account a probable moderating effect of CSR on the effect of legal enforcement of EM. They find that firms with high CSR tend to substitute REM through AEM in favour of the long-term objectives of shareholders and at the risk of a more likely detection of AEM. Their results show that in a strong legal enforcement, CSR firms tend to use even lower long-term adverse REM and higher AEM, which is easier for regulators to investigate. In short, CSR firms tend to emphasize share- and stakeholder value more than reporting perception or the probable enforcement consequences. All of these studies provide confirmation of the "myopia avoidance hypothesis". CSR firms act socially responsible due to their intrinsic motivation and are willing to trim back their performance goals to meet this incentive.

In a resumption of CSR incentive analysis, Porter and Kramer (2006) postulate that companies can increase their market value by acting socially responsible. This holds because they

strengthen their corporate image and reputation. This resumption introduces a new dimension of CSR incentives since entities without all types of responsible intentions may make use of CSR to gain a strategic positioning in the market. Several studies review a positive coherence between CSR performance and reputation enhancement (e.g., Verschoor 2005; Linthicum et al. 2010). The predominant CSR approaches can even enhance competitive advantages (e.g., Cohen 2009). If this holds true, then firms with low intrinsic motivation for social responsibility can also have incentives to invest in CSR activities, and it may be interesting to study their EM behaviour separately.

In the following, we elaborate a differentiation of the two groups of higher intrinsic motivation and higher reputational motivation to invest in CSR. Therefore, we make use of common voluntary disclosure considerations. With respect to CSR, reporting is one the most important factors in regard to signalling the corporate image to the market. Not exclusively but particularly for firms with a higher reputational motivation, the importance of an optimized CSR disclosure may therefore be of interest. Correspondingly, Graham et al. (2005) document in their survey of 400 USA firms that 92% of all CFOs regard enhancement of reputation as one major advantage of voluntary disclosure. In detail, voluntary CSR disclosure can function as a signal of the unobservable social responsibility of the firm to the market (e.g., Teoh & Hwang 1991). Mutual high CSR performance and optimized reporting can therefore help strengthen the corporate image of social responsibility. However, disclosed information requires certain characteristics to be able to influence decision-making processes. The most important among these characteristics are the relevance for the decisions of addressees and the reliability of the information (e.g., Dechow et al. 2010). Given the sophistication of society concerning CSR noted above, the presented information in CSR reports should be relevant to a variety of stakeholders. The reliability of the presented information is attached to the uncertainty of the addressees and the trustworthiness of the company. The use of commonly accepted reporting standards in the provision of CSR reports can be a necessary method of enhancing the reliability of the presented information. Naturally, all companies with vast investments in CSR may be interested in reporting their actions. However, they

may compare the benefits from a higher reliability of CSR information with the incurred costs for CSR reporting (e.g., Lev 1992).[128] Elaborate CSR reports that follow sustainability reporting guidelines incur more costs than CSR reports without this feature. Hence, the respective firm is likely to expect higher benefits from a disclosure; thus, higher costs to increase the reliability of the information are justifiable. Given these theoretical considerations, we assume that companies with high CSR performance and more reputational incentives (hereinafter, "high-reputation firms") will predominantly appear in the group with more costly CSR reports, whereas their more intrinsically motivated counterparts (hereinafter, "intrinsic firms") may be less distressed to invest more money in the perception of the CSR report.[129]

When thinking about their respective EM decisions, firms will compare the impact of the two EM types on the long-term value of the firm and the respective probability of accounting controversies, e.g., the detection of EM actions. Following these considerations, high-reputation firms may be interested in decreasing AEM to ensure that it does not negatively affect their reputation enhancement. Accounting controversies cause severe reputational damage and may unmask a firm's CSR incentive. A scaling back of AEM will likely prevent this from occurring. Meanwhile, intrinsic firms may act similarly due to their motivation to support stakeholders' decisions. However, high-reputation firms draw a more extrinsic picture than the frequently noted "myopia avoidance hypothesis". The prior literature on CSR and AEM has mostly ignored this incentive. Only KPW incorporate listing in *Fortune's America's Most Admired* companies as a control variable. Consideration of the listing may capture similar incentives compared to the disclosure of CSR reports and therefore shape reputation. However, they omit further considerations. We try to incorporate this dimension throughout our analyses to disentangle the various incentive groups, proposing the following hypothesis:

[128] Following the commonly known unraveling principle (e.g., Verrecchia 1983), CSR disclosure will only exist if the benefits exceed the costs. Thus, all companies providing a report assign a certain value to the reporting of their CSR activities.
[129] We cannot observe the actual incentive or motivation behind CSR investments, which is why we deliberately use the restriction "higher". It is likely that all firms investing in CSR have a certain degree of intrinsic motivation, which does not cancel out the use of CSR to influence corporate image and reputation. However, we propose that our differentiation captures the main incentives of the sample firms.

H_1: Ceteris paribus, firms with higher engagement in CSR use lower AEM. This coherence holds for both high-reputation and intrinsic firms.

An alternative explanation for a negative coefficient of CSR on EM is the positive association between firm performance and CSR. One major strand of the literature covers the coherence between CSR and firm performance, finding more support for a positive relationship (e.g., Griffin & Mahon 1997; Waddock & Graves 1997; Callan & Thomas 2009). Companies with a high performance may also be less distressed to take action in EM to improve their performance. We incorporate respective control variables to capture the influence of performance on our results.[130] The related literature additionally claims that CSR practices can be a means of opportunistic use to cover up selfish actions through CSR activities. In terms of financial reporting, selfish actions mean higher EM to mislead the addressees of financial reports in their decisions (e.g., Jensen & Meckling 1976; Davidson et al. 2004; Hemingway & Maclagan 2004; McWilliams et al. 2006; Petrovits 2006; DeMaCarty 2009). However, the respective studies have very mixed outcomes and concentrate on AEM. Moreover, increasing amounts of AEM will lead to a higher probability of accounting controversies, ultimately unmasking opportunistic CSR.[131] We claim that our more differentiated view of CSR can shape the incentives of firms to a higher degree, also because we apply differentiation between the use of AEM and REM in the following.

4.2.2 The earnings management trade-off relationship and the role of corporate governance

We study the two incentives for CSR investments in more detail by disentangling their influence on the use of REM and the trade-off decision between AEM and REM. The prior literature has shown that firms substitute between the two types of EM when considering their

[130] We also apply non-stationary cross-sectional regressions to compute our AEM and REM proxies to receive much more conservative discretionary proxies. We additionally incorporate an AEM model that cancels out the respective influences of firm performance on the magnitude of discretionary accruals to a higher degree than alternative AEM models. See sections 4.3.2.2 and 2.2.2.
[131] Although there are also results on a positive coefficient of CSR on EM, most of them are ambiguous for various EM measures (e.g., Trébucq & Russ 2005; Chih et al. 2008; Salewski & Zuelch 2014) or valid only in regulated industries (e.g., Prior et al. 2008; Kim & Venkatachalam, 2011).

costs and/or regulative conditions. Cohen et al. (2008) find that firms are aware of the possible costs associated with accounting investigations following AEM. On the other hand, Roychowdhury (2006), Cohen and Zarowin (2010) and Zang (2012) elaborate that REM requires alterations in business activities that can have severe impact on the future performance of the firm. Various factors shape the direction and intensity of this trade-off relationship. In particular, high enforcement, high monitoring through CG, auditing or regulative entities may limit the possibilities of using AEM without detection and may lead to a higher use of REM.

> "While auditors can second-guess the firm's accounting policies, they cannot readily challenge real economic actions […]." (Graham et al. 2005)

In their survey, Graham et al. (2005) confirm that CFOs use REM not only when needed in excess of AEM but as a substitute, despite knowing that they burn real future cash. The demand for a substitution is driven by the probability of being the subject of investigations, which increases with the amount of AEM used. Intrinsic firms may have a different evaluation of this situation since they supposedly weigh shareholder value more than a probable investigation, regardless of the fact that they will already use less EM. For high-reputation firms, accounting controversies may be a worst-case scenario and heavily damage their reputation. Hence, such an unmasking will burn all of the invested money in CSR engagement and reporting. Consequently, high-reputation firms will be more likely to use more REM than AEM at the expense of future shareholder value. Altogether, we propose the following hypothesis:

$\boldsymbol{H_2}$**:** Ceteris paribus, intrinsic firms substitute REM through AEM. This coherence reverses for high-reputation firms.

Moreover, CFOs are aware of the fact that the regulative and monitoring setting drives the probability of investigations concerning AEM (e.g., Chun et al. 2012). The importance of the company's board in monitoring the management properly is commonly known, particularly when ownership is widespread. Jensen (1993) notes that the board chair has the duty to control the status of the firm and represent the interests of shareholders responsibly. The board

places great importance on ensuring the presented picture reflects the financial situation to a high degree (e.g., Fama & Jensen 1983).

However, the related literature has mainly ignored considering the interplay of CSR strategies and CG as the monitoring device of the firm. There are only a few studies that incorporate CG when studying the coherence of CSR and EM, and none of them includes the trade-off relationship between REM and AEM. Choi et al. (2013) study ownership concentration in Korea and show that highly concentrated firms use higher AEM than other firms with the same levels of CSR. Cho and Chun (2016) analyse the interaction effects of various CG aspects on the relationship of CSR and REM using a composite CG index. KPW alternatively control for CG influences using an aggregate performance score, whereas BFMM completely omit CG aspects and focus on enforcement.[132]

We incorporate the effects of potentially non-linear CG aspects in the form of board independence on EM. Numerous studies on CG aspects show that independent members on the board are predestined to have an impartial view of the business and represent stakeholders' interests in their monitoring duty in a superior manner (e.g., Beasley 1996; Dechow et al. 1996; Xie et al. 2003; Davidson et al. 2005; Peasnell et al. 2005). Consequently, these studies show that a higher number of independent members on the board decreases the amounts of AEM used.[133] We use board independence and study its effect on both types of EM. In particular, it may be interesting to determine whether higher monitoring comes along with a higher or lower use of REM. Moreover, there may be a U-shaped relationship. Higher independence in the board may not infinitely improve monitoring. Rather, at a certain point, more members who are independent may decrease the monitoring ability of the board again since coordination and processing problems or a lack of expertise overcompensate for the independence effects (e.g., Jensen 1993; Bradbury et al. 2006). Following Haans et al. (2016), these two opposing effects may lead to a U-shaped or inverted U-shaped relationship between

[132] Naturally, an aggregate score or index can only control for the sum of strengths and weaknesses, and it might be useful to study the individual effect of certain monitoring devices.
[133] In addition, there is some evidence of a negative relationship between the proportion of outside directors and financial fraud (e.g., Beasley 1996; Dechow et al. 1996).

our independent and dependent variable, building a solid foundation for further considerations (e.g., Sutton & Staw 1995).[134]

We further incorporate the CSR incentives as a probable moderator of the effectiveness of CG aspects in limiting EM. As stated above, we distinguish between intrinsic and high-reputation firms. These incentives may shape the effects of monitoring, e.g., good CG, on both types of EM since they provide a foundation for the preference of AEM or REM as EM vehicles. Although intrinsic firms prefer AEM over REM, they may not be able to trade off from REM to AEM when higher monitoring limits AEM to a certain degree. In this case, there can be a counterbalancing effect on AEM, and we expect that CG works well in limiting EM. Conversely, high-reputation firms prefer REM over AEM, which is why they will use even more REM than AEM when AEM is limited, potentially leaving the magnitude of total EM unchanged. The use of REM will then increase with increasing independence in the board.[135] Altogether, we propose the following hypothesis:

$\boldsymbol{H_3}$**:** Ceteris paribus, CSR incentives moderate the relationship between CG and EM.

4.3 Research design

4.3.1 Data and sample selection

We use an initial dataset containing 10,637 firm-year observations from all European ASSET4 firms available in the Thomson Reuters database from 2005 to 2014. We exclude 1,277 observations in which disclosures following either the IFRS or US-GAAP standards were not made.[136] Financial institutions (SIC codes 6000-6999) account for 2,002 firm-years and are omitted due to the disparate balance sheet structure and business model, which leads to inaccurate EM measures. Furthermore, we require each firm-year grouping to consist of at

[134] Possible explanations of why firms exceed the optimal degree of independent members can be irrational decision-making, escalated commitment to CG or ambiguous goals (e.g., Haans et al. 2016).
[135] We are aware of the fact that this only holds if higher independence has no significant decreasing effect on REM, which is ex-ante undetermined.
[136] Although there are some differences in accounting standards between the IFRS and US-GAAP, they are highly aligned regarding basic principles, objectives and possibilities to pursue EM. Further, Van der Meulen et al. (2007) find no significant differences in the quality of accruals and the attributes of earnings for the two accounting systems, which is why we accept alignment.

least ten observations, which accounts for a loss of 939 observations. Subsequently, the paper uses an unbalanced panel with adequate information to conduct the analyses using AEM and REM proxies with regard to all controls and variables of interest. A total of 2.733 firm-year observations remain.

4.3.2 Measurement of applied variables

4.3.2.1 Measurement of CSR performance, CG characteristics and differentiation of CSR incentives

Numerous approaches to measuring CSR performance exist. KPW apply the extensively used (e.g., Turban & Greening 1997; Waddock & Graves 1997; Szwajkowski & Figlewicz 1999) measure proposed by KLD Research and Analytics Inc. (2006). It evaluates each firm on the strengths and weaknesses of several dimensions. They also construct the Social Index Domini 400, whose affiliation can be used as an alternative measure for CSR (e.g., McWilliams & Siegel 2000; Kim et al. 2012). In their international study, BFMM use an equally weighted CSR measure from EIRIS scaled between 3 and 12. We apply an overall score for the CSR performance of European companies using ASSET4 ESG data from Thomson Reuters. The data consist of 250 key performance indicators, classified within 18 categories or 4 pillars, for 967 European companies. The pillars are (1) environmental, (2) social, (3) CG and (4) economic performance. For each pillar, there are indicator z-scores that can be applied to calculate the pillar z-scores. Subsequently, we comprehend an overall CSR score assuming equal weights between pillars. For purposes of distinction in the further analyses, we exclude the CG pillar from the CSR score computation.[137] Instead, we use only the (1) environmental and (2) social performance to acquire an overall score (CSR_Score_{it}) for firm i in year t between 1 and 100.[138] Chatterji et al. (2016) investigate all major CSR ratings and document the high validity and agreement of ratings for the respective rater's location.

[137] Larcker et al. (2007) show that corporate governance importance correlates with separation of ownership and control and improves the extent to which managers act in line with stakeholders' interests. On the other hand, KPW claim that CSR incorporates social and environmental commitment. A distinct treatment and formulation of the two concepts is therefore mandatory, particularly since the relationship between corporate governance and CSR is not predetermined.

[138] To obtain an overall score between zero and one for better coefficient interpretation, we divide each score by 100. For further details about the CSR and CG performance scores, see Appendix E.

With KLD being the valid standard for CSR ratings within the USA, we consider ASSET4 ESG to be the valid approach for European analyses (e.g., Cheng et al. 2014; Huang & Watson 2015).

To distinguish high-reputation firms from intrinsic firms, we make use of the elaborated expected differences in CSR reporting. As stated above, the provision of a CSR report following internationally accepted guidelines for sustainable reporting can be viewed as a method of further increasing the reliability of the CSR reports and, therefore, the reliability and reputation of the firm. Therefore, we apply a reporting indicator (REP_{it}), which takes the value of one if firms collocate a CSR report and use GRI standards, zero otherwise.[139] They represent the global best practice approach for CSR reporting (e.g., Menichini & Rosati 2014; Vukic 2015; Calabrese et al. 2016; Lozano et al. 2016; Global Reporting Initiative 2017), optimize the presentation of CSR information and therefore maximize the probable outcomes concerning the reputation of the firm. Importantly, for our sample, there is no legal requirement to collocate any CSR report or to obey any reporting standards.[140] Using this indicator to separate the incentives, we define "intrinsic firms" as firms with above median net CSR performance scores, whereas "high-reputation firms" are firms with above median net CSR performance scores and G4 GRI CSR reporting. Although incentives will most likely only exist when considerable (above median) CSR investments are made, we will assess the validity of this definition in our uni- and multivariate analyses.

To control for CG aspects when studying CSR, we use the excluded, equally weighted CG pillar score (GOV_Score_{it}). When studying the moderating role of CSR in the relationship between CG and EM, we focus on explicit board independence, not the overall CG pillar. Therefore, we use the indicator value for board independence (BI_{it}) as the percentage of

[139] To ensure the validity of this indicator in the formulation, we test all ESG ASSET4 CSR reporting indicator values together with our new indicator in principal and maximum likelihood factor analyses to obtain the optimal representation of the variation of all factors by the combined factor used in the regressions. Untabulated results ensure the high validity of our proxy. The divergence in the computation of score-based CSR performance and indicator value-based CSR reporting dimension further demonstrates the validity of distinct interpretation.

[140] Our data still contain only observations where reporting is voluntary; meanwhile, the European Union has passed the Directive 2014/95/EU to oblige all companies of public interest to collocate a mandatory CSR report for all firm-years beginning with 2017.

strictly independent members on the total board.[141] For our multivariate analyses, we adjust the CG score to capture all aspects apart from the analysed independence characteristic. Therefore, we exclude the particular KPI category of board independence from the respective CG score calculation.[142]

4.3.2.2 Measurement of earnings management

As discussed in chapter 2.2, we use the cross-sectional McNichols (2002) model to comprehend AEM, while we apply the cross-sectional models for REM proposed by Roychowdhury (2006) to comprehend a combined signed measure REM_{it}, calculated in accordance with KPW for reasons of better comparison.[143]

Alternatively, we apply the two decile-based proxies for EM proposed by BFMM. Therefore, we sort $DACC\_abs_{it}$ and REM_{it} into deciles. We calculate a total EM proxy (EM_ALL_{it}) as the sum of the two deciles, whereas $REMvsAEM_{it}$ is a distinct proxy for the trade-off relationship between the two types of EM, calculated as the ratio of the REM_{it} decile and EM_ALL_{it}.

[141] We conform to the related literature and define independence as the percentage of members on the board who are not employed by the firm, have not served on the board for more than ten years, have holdings of less than 5 percent, do not have cross-board membership, do not have recent, immediate family ties to the firm and do not accept any compensation other than compensation for board service (e.g., Bradbury et al. 2006). See Appendix D for detailed definitions of the variables.
[142] See Appendix D for a detailed definition of these variables.
[143] $REM_{it} = CFO\_abn_{it} - PROD\_abn_{it} + DISEXP\_abn_{it}$. This proxy for real activities EM is negatively scaled.

4.3.3 Empirical models

To perform tests on the relationship between CSR performance and EM, we estimate the following basic empirical models:

Equation 26

$$
\begin{aligned}
DACC\_abs_{it}/TA_{it-1}\ (REM_{it}/TA_{it-1}) &= \alpha_0 + \beta_1 CSR\_Score_{it} + \beta_2 REP_{it} + \beta_3 REP_{it} * CSR\_Score_{it} \\
&+ \beta_4\, REM_{it}/TA_{it-1}\,(DACC\_abs_{it}/TA_{it-1}) + \beta_5 Size_{it-1} + \beta_6 MB_{it-1} \\
&+ \beta_7 LOSS_{it} + \beta_8 Leverage_{it-1} + \beta_9\, \Delta EMPLOY_{it}/EMPLOY_{it-1} \\
&+ \beta_{10}\, \Delta TA_{it}/TA_{it-1} + \beta_{11} GOV\_Score_{it} + \beta_{12}\, ACC_{it-1}/TA_{it-1} \\
&+ \beta_{13}\, CFO_{it}/TA_{it-1} + \beta_{14} ATO_{it-1} + \beta_{15} Litigation_{it} \\
&+ \beta_{16}\, \Delta GDP_{jt}/GDP_{jt-1} + \beta_{17} FirmAge_{it} + \beta_{18} Risk_{it} + \varepsilon_{it}
\end{aligned}
$$

Equation 27

$$
\begin{aligned}
EM\_ALL_{it}(REMvsAEM_{it}) &= \alpha_0 + \beta_1 CSR\_Score_{it} + \beta_2 REP_{it} + \beta_3 REP_{it} * CSR\_Score_{it} + \beta_4 Size_{it-1} \\
&+ \beta_5 MB_{it-1} + \beta_6 LOSS_{it} + \beta_7 Leverage_{it-1} \\
&+ \beta_8\, \Delta EMPLOY_{it}/EMPLOY_{it-1} + \beta_9\, \Delta TA_{it}/TA_{it-1} + \beta_{10} GOV\_Score_{it} \\
&+ \beta_{11}\, ACC_{it-1}/TA_{it-1} + \beta_{12}\, CFO_{it}/TA_{it-1} + \beta_{13} ATO_{it-1} \\
&+ \beta_{14} Litigation_{it} + \beta_{15}\, \Delta GDP_{jt}/GDP_{jt-1} + \beta_{16} FirmAge_{it} + \beta_{17} Risk_{it} \\
&+ \varepsilon_{it}
\end{aligned}
$$

CSR_Score_{it} is the continuous CSR performance score as stated in Appendix E. REP_{it} is the CSR reporting indicator discussed above, whereas $REP_{it} * CSR\_Score_{it}$ covers additional CSR performance effects for these firms. We use this semi-continuous formulation to assess the validity of our incentive definition. Due to the probable trade-off relationships between the two types of EM (e.g., Cohen et al. 2008; Zang 2012), we include either REM_{it} or $DACC\_abs_{it}$ as controls in **Equation 26**. They capture the substitution effects between EM

actions according to their relative costs.[144] Alternatively, we apply the two decile-based proxies EM_ALL_{it} or $REMvsAEM_{it}$ as the dependent variable in **Equation 27**, as a result of which separate trade-off controls are no longer required.

The incorporated control variables minimize omitted variable correlation biases, which would heavily constrain the validity of our analysis. Roychowdhury (2006) argues that the firm-specific size in terms of its market value of equity (MVE) and growth opportunities can explain considerable amounts of variation within EM proxies (e.g., Watts & Zimmerman 1986; Frankel et al. 2002; Prior et al. 2008; Ali & Zhang 2015). We include $Size_{it-1}$ as the lagged value of the natural logarithm of MVE. With regard to growth opportunities, we use MB_{it-1} as the ratio of the lagged market to book value of equity. A considerably high market to book ratio entails high expectations of future company growth, which EM can help meet. The two controls also stand for the significance of the equity incentives for EM, e.g., maximizing stock-based compensation (e.g., Cheng & Warfield 2005), since MB_{it}, MVE and stock-based compensation are similarly responsive to stock price changes.[145] $LOSS_{it}$ is an indicator that captures firms with negative earnings. They have different incentives and/or possibilities to pursue EM (e.g., Roychowdhury 2006). Kim and Park (2005) show that the degree of leverage ($Leverage_{it-1}$) can play an important role since highly leveraged companies may undergo contractual renegotiations. This can tempt them to increase income-decreasing EM. Leverage also embodies the influences of past performance. Following Ali and Zhang (2015), we include two measures for business growth. Total asset growth (ΔTA_{it}) captures the influences on the magnitude of accruals due to correlation with a firm's working capital. As the change in the number of employees, $\Delta EMPLOY_{it}$ contains a favourable growth and performance control (e.g., Zhang 2007). As discussed above, we use an aggregated CG score (GOV_Score_{it}) as a separate control for the influence of CG features on EM (e.g., Kim

[144] Following KPW, we alternatively use positive (negative) discretionary accruals as a dependent variable to estimate **Equation 26**. Moreover, we apply all distinct proxies of real-activities manipulation as a dependent variable to estimate **Equation 26**. We discuss the results in section 4.5.

[145] Prior research on EM has additionally claimed that the Big 4 auditors constrain EM better than other auditors (e.g., DeFond & Jiambalvo 1993; Becker et al. 1998; Francis et al. 1999). The Big 4 companies audit 97 percent of our sample, leaving almost no reference group influence, which is why we omit this control variable in our main analyses. However, we include it in our robustness analyses in section 4.5.

et al. 2012).[146] ACC_{it-1} and CFO_{it} control for influences on the variation of EM that are due to correlations with last year's accruals and prediction imperfection of discretionary accruals caused by the applied models (e.g., Young 1999; Ashbaugh et al. 2003).[147] Moreover, we implement ATO_{it} to control for the influence of a firm's positioning within the market. We use the decomposition of the return on net operating assets ($RONA$) into the profit margin (PM) and asset turnover (ATO) as well as the Du Pont model, which is broadly used in financial statement analysis (e.g., Lundholm & Sloan 2004; Palepu et al. 2004; Stickney et al. 2004; Penman 2009). ATO, the cost efficiency and scalability part of the decomposition, is the level of revenues divided by the net operating assets of the firm. $RONA$ is comprehended as the product of the two decomposition factors. Penman (2009) shows that a firm's and industry's characteristics concerning these two drivers are negatively correlated. The idea is that not only REM in particular (e.g., overproduction, sales manipulation) but also AEM are easier to pursue when firms bear higher ATO due to a higher scalability of operating assets leading to more transactions.[148] $Litigation_{it}$ controls for the influences of litigation risk, e.g., when regulation occurs (e.g., Bowen et al. 1995; Cheng & Warfield 2005) or auditors tend to emphasize conservative accounting (e.g., DeFond & Subramanyam 1998). ΔGDP_{jt} controls for macroeconomic effects in the respective state (e.g., Cohen et al. 2008). As a separate control for the influence of systematic market risk on EM, the beta factor of the respective firm ($Risk_{it}$) is applied. Prior research finds that firms with a higher beta engage in more EM (e.g., Warfield et al. 1995). Finally, to capture the effect of changing financial reporting behaviour due to maturation over time, we include the age of the firm ($FirmAge_{it}$; e.g., Kim et al. 2012).

[146] See Appendices D and E for the definition and calculation summary.
[147] The applied first-step models already include regressors to limit the existence of measurement errors. Nevertheless, the variables for growth, leverage and size control supplement a favourable specification that prevents the observed coherences between the variables of interest and the dependent variables from being spurious, which would lead to erroneous inferences (e.g., Kang & Sivaramakrishnan, 1995).
[148] In an earlier study, Jansen et al. (2012) detect EM scenarios using the divergence of PM and ATO. This is a first indicator that there is a connection between ATO and discretionary actions.

To study CSR incentives as a moderator for the effect of CG on EM, we use and extend **Equation 26** and **Equation 27** as follows. We add the linear and squared terms (BI_{it} and BI_{it}^2)[149] of board independence to formulate the relationship with EM. We capture the moderating effects using the interaction of board independence with a CSR incentive indicator for both intrinsic ($BI_{it} * CSR_{it}$) and reputation ($BI_{it} * CSR_{it} * REP_{it}$) firms.[150]

We use a FGLS Driscoll-Kraay estimator, which is robust to heteroscedasticity, autocorrelation, time dependence and cross-sectional (spatial) correlation (e.g., Driscoll & Kraay 1998). Simple variation and autocorrelation tests (e.g., Cameron & Trivedi 2010) show predominant between variation, heteroscedasticity and time series equi-correlation. De Hoyos and Sarafidis (2006) claim that considerable cross-sectional dependence can lead to inefficient fixed- and random-effects estimation with biased standard errors. Common robust (Arellano 1987; Froot 1989) and cluster robust (Rogers 1993) standard errors are inapplicable since they omit correlation across groups of cross-sections. However, cross-sectional dependence is determined by various factors, e.g., not only the correlation across multi-level cross-sections but also its complexion. In the case of endogeneity, standard fixed- and random-effects estimators become inconsistent, biased and inefficient. Instrumental variable (IV) fixed- or random-effects estimators with exogenous and relevant instruments remain. Considerable amounts of cross-sectional correlation lead to inconsistent and biased IV and Generalized Method of Moments estimators, which is a delicate issue. This applies for small panels with large-N asymptotics in linear-panel data models.[151] Testing for cross-sectional dependence is an issue within our empirical accounting data setting since the mutual existence of both cross-sectional correlation and endogeneity evokes a trade-off decision concerning these problems. We run Pesaran's CD test for unbalanced panels[152] and find a significant CD test

[149] We run simple curve estimations for the coherence of board independence and EM proxy to obtain a better understanding of which formulation fits the model best. We decide on a linear or non-linear (squared) formulation based on the level of significance, course of the curve and r-squared. The results favour a non-linear formulation and remain untabulated.
[150] We switch to an indicator based separation of our incentives following the definition above to be able to study the continuous, probably non-linear effects of board independence. Therefore, CSR_{it} is an indicator that captures all firms with above median CSR performance scores. The accuracy of this formulation is guaranteed when our CSR incentives are attached to above median CSR performances, which we analyze in section 4.4.2.1. However, we use alternative formulations in section 4.5.3.
[151] In particular, GMM estimators' large N asymptotics rely on the assumption of no cross-sectional correlation.
[152] Pesaran's CD test (Pesaran 2004) is applicable for small (N > T) unbalanced panels and is the suitable alternative to the common Lagrange multiplier (LM) test for long panels by Breusch and Pagan (1980).

statistic (0.01 levels), implying that cross-sectional correlation is an issue within the sample. Therefore, we use the specially designed Driscoll-Kraay estimator, which implements a Newey-West-type method to correct the cross-sectional averages of the moment conditions. It produces consistent and efficient variance-covariance matrices for the setting of heteroscedasticity, auto-correlation, time dependence and cross-sectional correlation with large-N, fixed T asymptotics.[153]

4.4 Results

4.4.1 Descriptive statistics and univariate analysis

Table 39 panel A and B present the sample distribution by two-digit SIC industries, country and year. The sample contains 20 different industries with a weighted average share of 6.93 percent and an average number of 137 observations. The total number of firm-years consistently increases from 2006 to 2012, and the highest degree is located in the United Kingdom (29.8 percent), followed by Germany (12.1 percent) and France (11.0 percent). Taking a closer look at the firms that use G4 GRI guidelines to collocate their CSR reports, we find that only 102 (8.80 percent) have below median CSR performances and that, in the group of firms with above median CSR score (high-CSR firms), 65.63 percent of the total sample disclose G4 GRI CSR reports and 34.37 percent either have none or do not invest in such an elaborate report. All continuous variables are winsorized at the top and bottom 1 percent of the respective distribution.

[153] The use of LSDV and FGLS Driscoll-Kraay estimators shows qualitatively equal results. When between variation is high compared to within variation, fixed effects are less efficient; thus, random effects estimators with dummies are used. Bootstrapped Hausman approves this estimation procedure. Hoechle (2007) notes that very small T decrease favourable estimator features again. We consider the panel $T = 8$ to be sufficiently large.

Table 39 – Sample description

Panel A: Distribution of firm-year observations by industry

Industry	Two-digit SIC	# of obs.	% of sample	Cum. %
Metal Mining	10	77	2.82%	2.82%
Oil and Gas Extraction	13	114	4.17%	6.99%
Building Construction General Contractor	15	102	3.73%	10.72%
Food and Kindred Products	20	157	5.74%	16.47%
Printing, Publishing, and Allied Industries	27	89	3.26%	19.72%
Chemicals and Allied Products	28	393	14.38%	34.10%
Stone, Clay, Glass, and Concrete Products	32	94	3.44%	37.54%
Primary Metal Industries	33	84	3.07%	40.61%
Industrial Machinery and Computer Equipment	35	237	8.67%	49.29%
Electronic and other Electrical Equipment	36	162	5.93%	55.21%
Transportation Equipment	37	155	5.67%	60.89%
Instruments and Related Products	38	144	5.27%	66.15%
Transportation by Air	45	10	0.37%	66.52%
Communications	48	230	8.42%	74.94%
Electric, Gas and Sanitary Services	49	130	4.76%	79.69%
Wholesale Trade-Durable Goods	50	57	2.09%	81.78%
Food Stores	54	110	4.02%	85.80%
Miscellaneous Retail	59	21	0.77%	86.57%
Business Services	73	238	8.71%	95.28%
Engineering and Management Services	87	129	4.72%	100.00%
		2,733	100.00%	

Panel B: Distribution of firm-year observations by country and year

	2006	2007	2008	2009	2010	2011	2012	2013	Total	(%)
Austria	5	5	7	7	7	7	6	8	52	(1.9)
Belgium	9	9	9	10	10	10	11	11	79	(2.9)
Cyprus	0	0	0	1	1	1	1	1	5	(0.2)
Denmark	11	13	13	13	13	14	14	14	105	(3.8)
Finland	5	5	7	7	9	9	9	9	60	(2.2)
France	27	32	35	37	41	43	43	43	301	(11.0)
Germany	29	33	40	45	45	46	46	46	330	(12.1)
Greece	1	0	0	1	1	1	2	2	8	(0.3)
Hungary	0	0	0	0	1	1	1	1	4	(0.2)
Ireland	5	5	5	6	7	7	7	7	49	(1.8)
Italy	17	19	21	21	22	23	24	24	171	(6.3)
Luxembourg	1	1	1	1	1	1	1	1	8	(0.3)
Netherlands	13	15	14	15	16	17	17	17	124	(4.5)
Norway	6	7	7	7	6	6	6	5	50	(1.8)
Poland	0	0	2	3	7	8	9	5	34	(1.2)
Portugal	3	3	3	3	3	3	3	3	24	(0.9)
Spain	5	5	7	7	9	10	12	13	68	(2.5)
Sweden	22	23	25	25	26	26	26	27	200	(7.3)
Switzerland	22	24	24	26	28	31	32	32	219	(8.0)
Turkey	0	0	2	2	4	4	6	10	28	(1.0)
United Kingdom	74	78	91	109	109	113	122	118	814	(29.8)
Total	**255**	**277**	**313**	**346**	**366**	**381**	**398**	**397**	**2,733**	**(100)**

Table 39 – Sample description (continued)

Panel C: Distribution of firm-year observations by CSR and REP			
High-CSR firms		**Low-CSR firms**	
1,574 (57.59%)		1,159 (42.41%)	
REP firms	**Non-REP firms**	**REP firms**	**Non-REP firms**
1,033 (65.63%)	541 (34.37%)	102 (8.80%)	1,057 (91.20%)
REP firms		**Non-REP firms**	
1,135 (41.53%)		1,598 (58.47%)	
2,733 (100%)			

A firm is defined as a high-CSR firm when the CSR score lies above the median of the full sample and defined as a low-CSR firm otherwise. A firm is defined as a REP firm when the firm discloses a CSR report in accordance with G4 GRI guidelines for sustainability reports and defined as a non-REP firm otherwise. The variables are defined in Appendix D.

Table 40 reports the descriptive statistics. The mean absolute value of discretionary accruals is 2.5 percent, whereas the mean value of discretionary accruals of zero indicates a desirable setting with no systematic up- or downward EM activity. On the other hand, absolute REM accounts for 17.9 percent, which is on average seven times the value of AEM. This type of EM seems to be more pronounced, which is in line with the findings of other studies (e.g., Graham et al. 2005; Cohen & Zarowin 2010; Zang 2012). The mean value of the signed real activities proxies of zero completes the favourable randomness setting.[154] Concerning the CSR performance score, the skewness of -0.61 signals an asymmetric distribution with a long tail to the left, indicating a disproportionate frequency of observations with very high CSR scores. The mean percentage of independent members on the board accounts for 23.6 percent on average. Concerning the control variables, the sample firms account for a low amount of losses (only 12 percent of all firm-year observations), the sizes and growth opportunities are high, and the mean leverage is 23.5 percent, with an interquartile range of 22.2 percent. Altogether, the sample characteristics are comparable to those of the prior literature. Several variable distributions bear high skewness and indicate the importance of outliers and winsorizing.

[154] Corresponding untabulated tests show that $DACC_{it}$ and REM_{it} are not significantly different from zero at the 0.01 level.

Table 40 – Descriptive statistics of relevant variables

Panel: Full sample	**N**	**Mean**	Q_{50}	**Std. dev.**	Q_{25}	Q_{75}
Dependent variables						
$DACC\_abs_{it}$	2,733	0.025	0.017	0.025	0.007	0.034
$DACC_{it}$	2,733	0.000	0.001	0.035	-0.015	0.017
REM_{it}	2,733	0.000	-0.011	0.251	-0.138	0.114
REM_abs_{it}	2,733	0.179	0.129	0.176	0.054	0.248
EM_ALL_{it}	2,733	10.997	11.000	4.057	8.000	14.000
$REMvsAEM_{it}$	2,733	0.502	0.500	0.208	0.364	0.667
Variables of interest						
CSR_Score_{it}	2,733	0.647	0.702	0.216	0.486	0.837
CSR_{it}	2,733	0.576	1.000	0.494	0.000	1.000
REP_{it}	2,733	0.415	0.000	0.493	0.000	1.000
BI_{it}	2,733	0.236	0.000	0.276	0.000	0.500
Control variables						
$Size_{it-1}$	2,733	15.282	15.196	1.663	14.033	16.436
MB_{it-1}	2,733	2.895	2.173	2.618	1.284	3.563
$LOSS_{it}$	2,733	0.120	0.000	0.325	0.000	0.000
$Leverage_{it-1}$	2,733	0.235	0.225	0.158	0.114	0.336
$\Delta EMPLOY_{it}$	2,733	0.050	0.025	0.179	-0.025	0.092
ΔTA_{it}	2,733	0.088	0.043	0.234	-0.022	0.134
GOV_Score_{it}	2,733	0.565	0.586	0.165	0.449	0.691
ACC_{it-1}	2,733	-0.048	-0.041	0.063	-0.073	-0.015
CFO_{it}	2,733	0.114	0.099	0.085	0.064	0.147
ATO_{it-1}	2,733	0.936	0.820	0.563	0.560	1.13
$Litigation_{it}$	2,733	0.212	0.000	0.409	0.000	0.000
ΔGDP_{jt}	2,733	0.029	0.043	0.086	-0.034	0.081
$FirmAge_{it}$	2,733	2.563	2.708	0.529	2.398	2.944
$Risk_{it}$	2,733	0.968	0.950	0.442	0.670	1.250

All variables are defined in Appendix D & E.

Table 41 considers the descriptive statistics comparing high- and low-CSR firms in panel A. The mean (median) CSR_Score_{it} for the high-CSR firms is 0.808 (0.822) and for the low-CSR firms is 0.429 (0.450). There are significant differences in the EM values (mean REM_{it} of 0.015 (-0.021) for high-CSR (Low-CSR) firms and the mean $DACC\_abs_{it}$ of 0.022 (0.028) for high-CSR (low-CSR) firms. Moreover, the aggregate EM measure (EM_ALL_{it}) significantly differs between high- and low-CSR firms, which underpins that high-CSR firms engage in lower EM than their socially less responsible counterparts. There are more high-CSR firms collocating a G4 GRI CSR report than low-CSR firms, in which the differential is significant again. However, we cannot find differences in the CG aspects of the two groups. In summary, these findings support that high-CSR firms, independent of their incentive, on average engage in less EM than firms with low CSR investments and, therefore, performance. Moreover, high-CSR firms are significantly larger, though they have higher leverage and higher systematic market risk than low-CSR firms. They entail lower actual growth

($\Delta EMPLOY_{it}$ and ΔTA_{it}) and expected growth opportunities (MB_{it-1}). High-CSR firms are on average older and most likely more consolidated. Panel B compares firms with G4 GRI CSR reporting to the remaining sample. The results on the differences in EM proxies show that firms with this elaborate, though costly, CSR reporting use considerably lower EM. However, the significance of the difference decreases compared to panel A, which could mean that firms with high CSR performance but without G4 GRI CSR reporting (intrinsic CSR firms) use considerably lower EM and therefore decrease the mean values of the EM proxies for the non-REP firms. Another reason could be that high-reputation firms use considerably more EM than the mean high-CSR firm. Both could prove that the separation of the two CSR incentives is valid. Moreover, CSR scores are significantly different and show that most firms with G4 GRI CSR reporting belong to the high-CSR firms group. The numbers in panel C of **Table 41** confirm this proposition. Naturally, firms with few CSR actions that potentially improve reputation will not invest more money in the provision of G4 GRI CSR reporting to cultivate their image. On the other hand, firms that spend the money may mutually have extensive CSR to maximize the possible outcomes for their reputation.

Table 41 – Descriptive statistics by high- and low-CSR firms and REP and non-REP firms

Panel A: Full sample

	High-CSR firms (1)			Low-CSR firms (2)			Difference (1) – (2): p-value	
	N	Mean	Q_{50}	N	Mean	Q_{50}	Wilcoxon test	t-test
Dependent variables								
$DACC\_abs_{it}$	1,574	0.022	0.015	1,159	0.028	0.019	<0.001	<0.001
REM_{it}	1,574	0.015	-0.004	1,159	-0.021	-0.028	<0.001	<0.001
EM_ALL_{it}	1,574	10.636	10.000	1,159	11.487	11.000	<0.001	<0.001
$REMvsAEM_{it}$	1,574	0.502	0.500	1,159	0.500	0.500	0.935	0.808
Variables of interest								
CSR_Score_{it}	1,574	0.808	0.822	1,159	0.429	0.450	<0.001	<0.001
REP_{it}	1,574	0.656	1.000	1,159	0.088	0.000	<0.001	<0.001
BI_{it}	1,574	0.239	0.000	1,159	0.233	0.000	0.948	0.549
Control variables								
$Size_{it-1}$	1,574	15.889	15.885	1,159	14.458	14.344	<0.001	<0.001
MB_{it-1}	1,574	2.645	2.012	1,159	3.234	2.498	<0.001	<0.001
$LOSS_{it}$	1,574	0.111	0.000	1,159	0.131	0.000	0.112	0.112
$Leverage_{it-1}$	1,574	0.248	0.234	1,159	0.216	0.201	<0.001	<0.001
$\Delta EMPLOY_{it}$	1,574	0.034	0.016	1,159	0.072	0.036	<0.001	<0.001
$\Delta ASSET_{it}$	1,574	0.065	0.036	1,159	0.121	0.056	<0.001	<0.001
GOV_Score_{it}	1,574	0.611	0.629	1,159	0.503	0.518	<0.001	<0.001
ATO_{it-1}	1,574	0.910	0.800	1,159	0.973	0.840	0.182	0.004
$FirmAge_{it}$	1,574	2.662	2.773	1,159	2.428	2.639	<0.001	<0.001
$Risk_{it}$	1,574	1.003	0.990	1,159	0.921	0.830	<0.001	<0.001

Panel B: Full sample

	REP firms (3)			Non-REP firms (4)			Difference (3) – (4): p-value	
	N	Mean	Q_{50}	N	Mean	Q_{50}	Wilcoxon test	t-test
Dependent variables								
$DACC\_abs_{it}$	1,135	0.023	0.016	1,598	0.025	0.017	0.210	0.053
REM_{it}	1,135	0.011	-0.007	1,598	-0.008	-0.014	0.126	0.059
EM_ALL_{it}	1,135	10.803	11.000	1,598	11.135	11.000	0.046	0.035
$REMvsAEM_{it}$	1,135	0.502	0.500	1,598	0.501	0.500	0.961	0.906
Variables of interest								
CSR_Score_{it}	1,135	0.803	0.837	1,598	0.536	0.535	<0.001	<0.001
BI_{it}	1,135	0.220	0.000	1,598	0.248	0.167	0.001	0.010
Control variables								
$Size_{it-1}$	1,135	16.140	16.166	1,598	14.672	14.579	<0.001	<0.001
MB_{it-1}	1,135	2.536	1.936	1,598	3.150	2.414	<0.001	<0.001
$LOSS_{it}$	1,135	0.113	0.000	1,598	0.125	0.000	0.351	0.351
$Leverage_{it-1}$	1,135	0.257	0.242	1,598	0.219	0.205	<0.001	<0.001
$\Delta EMPLOY_{it}$	1,135	0.035	0.013	1,598	0.062	0.032	<0.001	<0.001
$\Delta ASSET_{it}$	1,135	0.060	0.030	1,598	0.109	0.055	<0.00	<0.001
GOV_Score_{it}	1,135	0.609	0.622	1,598	0.534	0.549	<0.001	<0.001
ATO_{it-1}	1,135	0.866	0.760	1,598	0.987	0.860	<0.001	<0.001
$FirmAge_{it}$	1,135	2.697	2.833	1,598	2.467	2.708	<0.001	<0.001
$Risk_{it}$	1,135	1.007	0.980	1,598	0.940	0.905	<0.001	<0.001

A firm is defined as a high-CSR firm when the CSR score lies above the median of the full sample and defined as a low-CSR firm otherwise. A firm is defined as a REP firm when the firm discloses a CSR report in accordance with G4 GRI guidelines for sustainability reports defined as a non-REP firm otherwise. We evaluate the significances of means and medians based on t-tests and Wilcoxon tests, respectively. We calculate the p-values and Z-statistics based on two-tailed tests. The variables are defined in Appendix D.

Table 42 reports the correlation matrix. CSR_Score_{it} and REP_{it} are negatively correlated with $DACC\_abs_{it}$ and EM_ALL_{it} but positively correlated with REM_{it}. Higher CSR firms seem to use less EM. The correlations for REP_{it} are qualitatively equal, though less pronounced. This could be another sign that these firms act similarly, although not equally, making differentiated analyses a meaningful objective.

Table 42 – Correlations among CSR score, EM proxies, and other selected variables

Correlations among CSR score, EM proxies, and other selected variables

		1	2	3	4	5	6	7	8	9
1.	$DACC\_abs_{it}$	1.000								
2.	REM_{it}	0.020	1.000							
3.	EM_ALL_{it}	0.592***	-0.626***	1.000						
4.	$REMvsAEM_{it}$	-0.520***	-0.640***	-0.021	1.000					
5.	CSR_Score_{it}	-0.109***	0.104***	-0.133***	-0.011	1.000				
6.	REP_{it}	-0.037*	0.036*	-0.040**	0.002	0.611***	1.000			
7.	BI_{it}	0.061***	-0.003	0.036*	-0.042**	0.057***	0.035*	1.000		
8.	$Size_{it-1}$	-0.017	0.139***	-0.100***	-0.091***	0.465***	0.435***	-0.034*	1.000	
9.	MB_{it-1}	0.061***	0.209***	-0.070***	-0.216***	-0.148***	-0.116***	-0.002	0.079***	1.000
10.	$LOSS_{it}$	0.256***	-0.081***	0.190***	-0.037*	-0.051***	-0.018	0.020	-0.177***	-0.161***
11.	$Leverage_{it-1}$	-0.069***	-0.072***	-0.017	0.110***	0.080***	0.120***	-0.066***	0.063***	-0.021
12.	$\Delta EMPLOY_{it}$	0.021	0.034*	-0.004	-0.054***	-0.112***	-0.075***	-0.003	0.021	0.175***
13.	$\Delta ASSET_{it}$	-0.010	0.076***	-0.049**	-0.079***	-0.123***	-0.103***	-0.029	0.007	0.231***
14.	GOV_Score_{it}	0.029	0.022	0.026	-0.049**	0.404***	0.224***	0.402***	0.066***	-0.074***
15.	ATO_{it-1}	-0.047**	-0.001	-0.018	0.035**	-0.046**	-0.106***	0.008	-0.177***	0.195***
16.	$FirmAge_{it}$	-0.087***	0.040**	-0.074***	0.029	0.256***	0.215***	0.013	0.125***	-0.174***
17.	$Risk_{it}$	0.065***	0.026	0.048**	-0.073***	0.093***	0.074***	0.047**	0.060***	-0.070***

Correlations among CSR score, EM proxies, and other selected variables (continued)

		10	11	12	13	14	15	16	17
10.	$LOSS_{it}$	1.000							
11.	$Leverage_{it-1}$	0.081***	1.000						
12.	$\Delta EMPLOY_{it}$	-0.136***	-0.121***	1.000					
13.	$\Delta ASSET_{it}$	-0.205***	-0.131***	0.588***	1.000				
14.	GOV_Score_{it}	-0.025	-0.048**	-0.044**	-0.048**	1.000			
15.	ATO_{it-1}	-0.120***	-0.257***	-0.003	0.031	0.008	1.000		
16.	$FirmAge_{it}$	0.009	0.014	-0.144***	-0.145***	0.108***	0.015	1.000	
17.	$Risk_{it}$	0.124***	0.087***	0.001	-0.005	0.083***	-0.128	0.053***	1.000

*, **, and *** indicate significance at the 0.10, 0.05 and 0.01 levels, respectively, based on two-tailed tests. The variables are defined in Appendix D.

4.4.2 Multivariate analysis

4.4.2.1 The influence of corporate social responsibility on earnings management

Table 43 shows the results for the multivariate regression analyses of CSR and EM based on **Equation 26** and **Equation 27**. The first two columns contain separate analyses of the two types of EM ($DACC\_abs_{it}$ and REM_{it}), whereas column three shows the outcomes for the total proxy for EM (EM_ALL_{it}). All regressions are highly significant ($p < 0.01$), and the explanatory power of the models varies from 23 percent for $DACC\_abs_{it}$ to 26 percent for REM_{it}. These figures are slightly higher than those reported by KPW and BFMM. Initially, our results concerning the influence of CSR engagement on AEM and REM seem to be highly consistent with these authors. The coefficient for CSR_Score_{it} is negative and significant at the 0.01 level in columns one and three, indicating that firms with a higher score-based performance in CSR engage in less AEM and total EM.[155] The coefficient on REM_{it} is positive and highly significant (at the 0.01 level), signifying a lower use of REM. In summary, these results can indicate that firms with higher CSR engagement engage less in EM than their less responsible counterparts.

When we take a closer look at the specific CSR incentives, the findings are much more diverse. The coefficient for firms with G4 GRI CSR reporting is significant, as is the isolated influence of this group's CSR score on AEM and total EM ($CSR\_Score_{it} * REP_{it}$). In summary, they suggest that the direction of the total effect for firms with this elaborate reporting depends on the respective CSR engagement. For example, for $DACC\_abs_{it}$, all high-reputation firms exceeding a net CSR performance score of 0.68 have a combined negative effect, indicating lower AEM activities. Thus, they decrease the risk of an accounting controversy that might negatively affect their image. With a mean CSR performance score of 0.80, most firms with G4 GRI CSR reporting (88 percent) and, more importantly, almost every (97 percent) high-reputation firm (firm with

[155] A consideration of signed discretionary accruals ($DACC_{it}$) shows the same results. We also use various alterations and divisions of the EM proxies in additional analyses in section 4.5.2. However, the confirmative results remain untabulated.

high CSR and G4 GRI CSR reporting) use even lower AEM than the remaining sample.[156] This result also confirms our formulation for the moderating role of the two CSR incentives in the influence of CG on EM.

Table 43 – Multiple regressions of EM on CSR

	$DACC\_abs_{it}$	REM_{it}	EM_ALL_{it}	$REMvsAEM_{it}$
CSR_Score_{it}	-0.004	0.225	-2.620	-0.120
	(3.83)***	(13.32)***	(12.02)***	(16.56)***
REP_{it}	0.023	0.046	2.325	-0.041
	(6.35)***	(0.92)	(11.09)***	(2.34)**
$CSR\_Score * REP_{it}$	-0.030	-0.093	-2.721	0.076
	(7.75)***	(1.81)*	(10.07)***	(3.60)***
REM_{it}	0.003			
	(2.66)***			
$DACC\_abs_{it}$		0.270		
		(2.90)***		
$Size_{it-1}$	0.001	0.014	-0.083	-0.010
	(3.10)***	(5.91)***	(2.44)**	(3.99)***
MB_{it-1}	0.000	0.013	-0.096	-0.009
	(0.45)	(5.19)***	(1.94)*	(7.41)***
$LOSS_{it}$	0.020	0.023	1.426	-0.075
	(24.72)***	(3.43)***	(17.30)***	(23.41)***
$Leverage_{it-1}$	-0.012	-0.096	-0.062	0.109
	(6.25)***	(6.56)***	(0.19)	(10.22)***
$\Delta EMPLOY_{it}$	0.001	-0.049	0.613	0.028
	(1.14)	(2.72)***	(4.29)***	(5.50)***
ΔTA_{it}	0.000	-0.022	0.311	-0.005
	(0.76)	(1.00)	(1.14)	(0.42)
GOV_Score_{it}	0.006	-0.056	1.502	0.002
	(3.20)***	(3.83)***	(6.13)***	(0.14)
ACC_{it-1}	-0.016	-0.260	2.597	0.188
	(6.29)***	(8.57)***	(5.73)***	(8.02)***
CFO_{it}	0.000	1.280	-14.383	-0.768
	(0.04)	(39.27)***	(31.26)***	(33.90)***
ATO_{it-1}	0.000	-0.036	0.610	0.034
	(0.77)	(13.10)***	(5.26)***	(15.78)***
$Litigation_{it}$	-0.000	0.157	-1.488	-0.066
	(0.19)	(16.60)***	(5.86)***	(7.63)***
ΔGDP_{jt}	0.000	-0.043	1.371	0.044
	(0.02)	(0.80)	(1.82)*	(2.00)**
$FirmAge_{it}$	-0.001	0.031	-0.302	-0.014
	(2.32)**	(5.88)***	(3.36)***	(9.22)***
$Risk_{it}$	0.002	0.028	0.163	-0.033
	(4.53)***	(6.26)***	(2.69)***	(10.25)***
Year Dummies	Included	Included	Included	Included
Industry Dummies	Included	Included	Included	Included
R^2	0.23	0.26	0.24	0.23
N	2,733	2,733	2,733	2,733

*, **, and *** indicate significance at the 0.10, 0.05, and 0.01 levels, respectively, based on two-tailed tests. All test statistics are calculated using Driscoll-Kraay standard errors with lag (7). The mean VIF for all regressions are again considerably below 5.

[156] The remaining 12 percent of the G4 GRI CSR reporting firms can have lower budgets for CSR investments and/or instead decide to use more EM to report higher reported earnings. These firms seem to miss the point of the higher risk of accounting controversies.

In column two, the coefficient for the interaction term $CSR\_Score_{it} * REP_{it}$ attracts attention. It is negative and significant at the 0.10 level[157], indicating a probable trade-off between AEM and REM for the G4 GRI CSR reporting firms. Compared to non-REP firms, they seem to engage in more REM with increasing CSR engagement and decreasing AEM, as noted above. Column 4 covers the trade-off proxy between REM and AEM and underpins this finding, which also fits the considerations in the prior literature concerning the trade-off between the two types of EM (e.g., Graham et al. 2005; Cohen et al. 2008; Badertscher 2011; Kim et al. 2012; Zang 2012). Here, the coefficients on REP_{it} and $CSR\_Score_{it} * REP_{it}$ show highly significant proof of a substitution of AEM through REM when the CSR score rises compared to the remaining firms, which tend to trade off vice versa. This is already the case for firms with G4 GRI CSR reports and a net CSR score higher than 0.56.[158] Consequently, high-reputation firms may be more willing to exchange future shareholder value for an improved image in the present. Lower AEM will decrease their risk of accounting investigations and, therefore, reputational damage. Higher REM will ensure that the company still provides sufficient earnings numbers, and business activities will need to be altered based on the risk of future performance. This could be viewed as a hint that reputation is instead a mix of earnings numbers and CSR perception and that the myopia hypothesis may not hold when firms use CSR to build their reputation rather than out of intrinsic motivation. Intrinsic firms seem to weigh future shareholder value much higher (e.g., Bozzolan et al. 2015) and therefore substitute from REM to AEM at a higher risk of accounting controversies. They may have no pronounced incentive to hide their EM activities to cultivate their image.[159] It is important to view the latter outcome in line with column 3, as addressed above. Intrinsic firms use a lower amount of EM overall, whereas high-reputation firms with a CSR score of below 0.85, i.e., 48 percent of all high-reputation firms, use a significantly higher total EM compared to the remaining firms.

[157] We explain this very weak coherence with the issue that negative values of REM stand for income-decreasing EM, while it is uncertain whether the discussed situations can actually be used into the opposite direction. In chapter 4.5, we alternatively use only positive values of REM and find much more robust results.
[158] There are no firms with CSR scores below 0.56 that are located in the high-reputation firms group (high-CSR firms with G4 GRI CSR report).
[159] On a final note, the existence of a trade-off relationship is confirmed by the two significant positive coefficients for REM_{it} ($DACC\_abs_{it}$) in column one (two).

Further, companies with a higher market to book ratio and, in particular, a higher market value of equity seem to exhibit more AEM than REM, most likely to maintain long-term profitability and to maximize shareholder value (e.g., Frankel et al. 2002; Roychowdhury 2006). Similarly, equity incentives to achieve stock-based compensation may arise late in the year, in which changes in business activities consume too much time (Cheng & Warfield 2005). Companies with higher leverage seem to use considerably lower AEM and higher REM, which can be due to comprehensive assessment by lenders, which makes it more difficult for the respective firms to engage in AEM (e.g., Becker et al. 1998; Ali & Zhang 2015). Concerning the introduced variable of ATO_{it-1}, the results show that firms with a higher ATO exhibit higher REM due to the higher scalability of business processes. The coefficient is highly significant (at the 0.01 level). Moreover, there seems to be another trade-off that decreases AEM and increases REM, which confirms the hypothesis of easier ways to perform REM when ATO is higher.

In summary, our analyses indicate that different CSR incentives exist and can lead to significant differences in EM decisions. Our findings for intrinsic firms are consistent with those of KPW, Hong and Andersen (2011) and BFMM, whereas for high-reputation firms, we find considerable deviations from the myopia hypothesis, supporting $\boldsymbol{H_1}$ and $\boldsymbol{H_2}$.

4.4.2.2 The influence of corporate governance on EM and the moderating role of CSR

Table 44 reports the results concerning multiple regressions of CSR and board independence (BI_{it}) on the EM proxies. The results on the influence of CSR on EM for the two separate incentives remain qualitatively almost equal. The coefficients for the net CSR score on REM_{it} ($DACC\_abs_{it}$, EM_ALL_{it} and $REMvsAEM_{it}$) are positive (negative) and significant. Together with the coefficients on REP_{it} and $CSR\_Score_{it} * REP_{it}$ they confirm our considerations from **Table 43**.

Table 44 – Multiple regressions of EM on CSR and BI

	$DACC\_abs_{it}$	REM_{it}	EM_ALL_{it}	$REMvsAEM_{it}$
CSR_Score_{it}	-0.003	0.219	-2.605	-0.113
	(1.77)*	(12.94)***	(11.03)***	(17.92)***
REP_{it}	0.023	0.037	2.415	-0.032
	(6.94)***	(0.77)	(10.90)***	(1.97)**
$CSR\_Score * REP_{it}$	-0.031	-0.076	-2.963	0.061
	(9.23)***	(1.46)	(10.16)***	(2.93)***
BI_{it}	-0.006	-0.113	0.334	0.179
	(1.79)*	(5.13)***	(2.24)**	(9.47)***
${BI^2}_{it}$	0.018	0.159		-0.290
	(3.95)***	(4.37)***		(10.94)***
$BI * CSR_{it}$	-0.005	-0.004	0.037	-0.001
	(3.34)***	(0.39)	(0.22)	(0.05)
$BI * CSR * REP_{it}$	0.002	-0.038	0.538	0.035
	(0.94)	(3.52)***	(2.44)**	(3.11)***
REM_{it}	0.003			
	(2.60)**			
$DACC\_abs_{it}$		0.261		
		(2.82)***		
$Size_{it-1}$	0.001	0.014	-0.084	-0.009
	(3.29)***	(5.89)***	(2.31)**	(4.64)***
MB_{it-1}	0.000	0.013	-0.100	-0.009
	(0.55)	(5.45)***	(2.04)**	(7.34)***
$LOSS_{it}$	0.020	0.020	1.457	-0.071
	(24.32)***	(3.22)***	(18.16)***	(21.22)***
$Leverage_{it-1}$	-0.012	-0.100	-0.006	0.113
	(6.32)***	(6.95)***	(0.02)	(11.18)***
$\Delta EMPLOY_{it}$	0.002	-0.050	0.635	0.027
	(1.35)	(2.81)***	(4.59)***	(5.34)***
ΔTA_{it}	0.000	-0.022	0.306	-0.005
	(0.77)	(1.02)	(1.14)	(0.48)
GOV_Score_{it}	0.004	-0.025	1.133	-0.014
	(1.62)	(2.25)**	(5.22)***	(0.87)
ACC_{it-1}	-0.016	-0.256	2.580	0.183
	(6.27)***	(8.37)***	(5.71)***	(7.79)***
CFO_{it}	-0.000	1.280	-14.368	-0.769
	(0.04)	(41.58)***	(32.60)***	(35.38)***
ATO_{it-1}	0.000	-0.036	0.618	0.034
	(1.02)	(11.78)***	(5.33)***	(16.40)***
$Litigation_{it}$	-0.000	0.159	-1.487	-0.070
	(0.08)	(17.01)***	(5.80)***	(7.27)***
ΔGDP_{jt}	-0.002	-0.047	1.317	0.057
	(0.90)	(1.01)	(1.81)*	(4.13)***
$FirmAge_{it}$	-0.001	0.032	-0.298	-0.016
	(1.93)*	(6.17)***	(3.40)***	(11.26)***
$Risk_{it}$	0.002	0.028	0.146	-0.033
	(4.45)***	(6.66)***	(2.54)**	(10.41)***
Year Dummies	Included	Included	Included	Included
Industry Dummies	Included	Included	Included	Included
R^2	0.23	0.26	0.24	0.24
N	2,733	2,733	2,733	2,733

*, **, and *** indicate significance at the 0.10, 0.05, and 0.01 levels, respectively, based on two-tailed tests. All test statistics are calculated with Driscoll-Kraay standard errors with lag (7). The mean VIF for all regressions are again below 5. BI_{it} is the respective percentage of strictly independent members on the board, and ${BI^2}_{it}$ is the respective squared term. See Appendix D for detailed variable descriptions. CSR_{it} is an indicator that captures high-CSR firm as firms with a CSR performance score above the median of the full sample.

The linear coefficient for board independence on $DACC\_abs_{it}$ is negative and significant, which is consistent with prior research using a solely linear formulation (e.g., Beasley 1996; Dechow et al. 1996; Klein 2002; Davidson et al. 2005; Peasnell et al. 2005). However, we find the proposed squared relationship with a turning point, from whence the effects of higher independence on EM reverse. We estimate the 95 percent confidence interval for the turning points of all of our curves using the Fieller method (e.g., Fieller 1954; Haans et al. 2016) to construct the interval. We find that all turning points are located well within these intervals and data range, which is why the non-linearities are in fact (inverted) U-shaped relationships.[160] Consequently, the course of the curve is U-shaped for AEM, e.g., the positive monitoring effect is gradually offset by the negative coordination/expertise effects.[161] The linear interaction effect $BI_{it} * CSR_{it}$ is highly significant (at the 0.01 level) and negative (coefficient of -0.005), indicating that high-CSR firms use significantly lower AEM than low-CSR firms as BI_{it} increases. This leads to different slopes of the U-shaped curves. Higher board independence decreases the use of AEM for (low-) high-CSR firms until the degree of (16.02) 29.37 percent independent members. At which point, the overall effect for (low-) high-CSR firms increases again and even becomes positive at the degree of (32.04) 58.74 percent board independence. The monitoring effect seems to be higher for high-CSR firms, which is why it takes longer for coordination/expertise problems to set them off. This could be due to the increased responsiveness of the CSR incentives to a limitation of AEM since both reputation and intrinsic incentives tend to decrease the usage of AEM in the first place due to the risk of accounting controversies and intrinsic motivation. Hereafter, unsurprisingly, there is no additional significant interaction effect for high-reputation firms. Regardless, the results show that excessive independence can also lead to the opposite effect on EM.

Figure 7 shows the course of the curve for the combined BI_{it} effects on AEM separately for the total sample and both the low- and high-CSR firms and confirms our finding that the effect of board independence on AEM is more pronounced for high-CSR firms.

[160] The tests of simple curve regressions and the course of the curve for the coherence of our EM and CG proxies confirm this finding. We additionally test for flattening/steepening of the curve through the moderator. Since our moderators are indicators, situations of flattening and steepening are less likely. Tests including the incorporation of squared moderation terms support this position (e.g., Haans et al., 2016).
[161] In this sense, a positive effect means decreasing the use of AEM, whereas a negative effect means increasing the use of AEM.

Figure 7 – The non-linear influence of BI on AEM

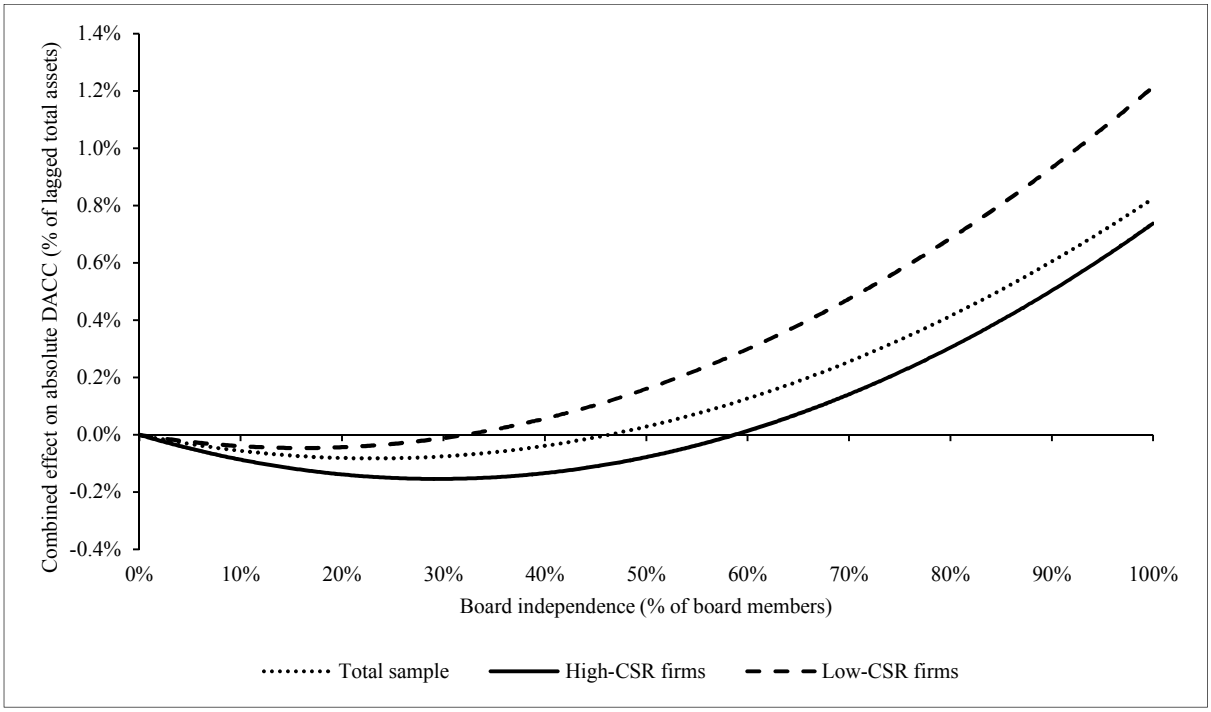


Figure illustrates the combined effect of the linear and squared coefficients for board independence on ***DACC_abs***. The respective lines show the course for the total sample (2,733 obs.), high-CSR firms (1,574 obs.) and low-CSR firms (1,159 obs.) separately.

Interestingly, the coefficient for BI_{it} in column three indicates that the total EM engagement increases (coefficient 0.055 at the 0.05 level).[162] In addition, we have a significant and even higher coefficient for the interaction effect of BI_{it} and the high-reputation firms group, signifying that these firms apply significantly more total EM than all other firms. The coefficients on REM in column two can explain this finding. First, we find a negative linear and a positive squared coefficient (both significant at the 0.01 level), suggesting another U-shaped relationship. Bearing in mind that REM is negatively scaled, a higher BI_{it} increases the amount of REM used. This result is interesting since we can interpret it as a probable shortcoming of board independence as a monitoring device in limiting REM. Moreover, high-reputation firms seem to use even more REM as the independence of the board increases, coinciding with our assumptions concerning how executives would act to maintain decent earnings numbers that cultivate their reputation strategy. **Figure 8** shows the course of the curve for the combined BI_{it} effect on REM, this time

[162] For board independence on total EM coherence, our analyses show that sole linear formulation fits best. This is the only regression in which squared formulation cannot help improve the fit of the coherence.

separately for the reputation and low-CSR firms as well as the total sample. Further calculations show that, for high-reputation firms, the turning point is located at a BI_{it} of 47.67 percent, with a zero at 95.34 percent, indicating that the effect increases the use of REM for most of the curve.

Figure 8 – The non-linear influence of BI on REM

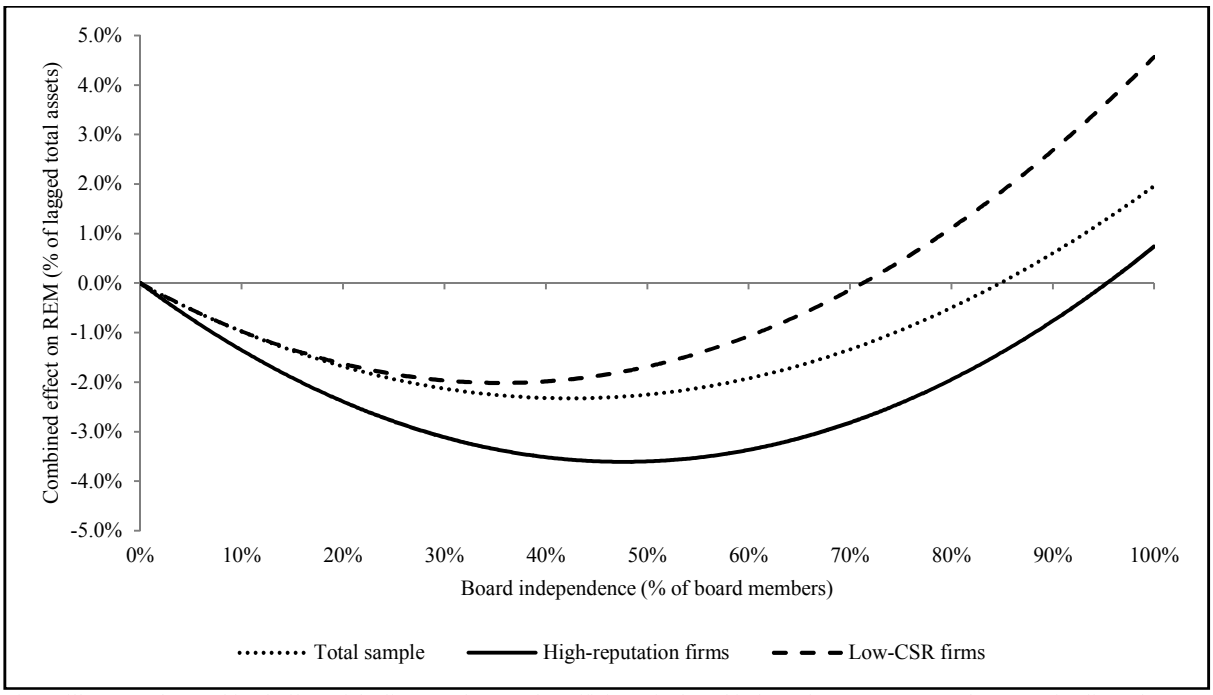


Figure illustrates the combined effect of the linear and squared coefficients for board independence on $\boldsymbol{REM_{it}}$. The respective lines show the course for the total sample (2,733 obs.), high-reputation firms (1,033 obs.) and low-CSR firms (1,159 obs.) separately.

Column four shows the results for the EM trade-off proxy. Now, we find an inverted U-shaped relationship, which is consistent with our previous results. Firms seem to substitute AEM through REM as monitoring increases until a certain degree of independence in which the monitoring effect is offset and hence possibilities for AEM increase again. The interaction effect for high-reputation firms is significant and shows that these firms may be more inclined to trade off from AEM to REM, which is in line with the other results and the elaborated incentives. For high-reputation firms, we comprehend a turning point at 36.90 percent, with a zero at 73.80 percent. At which point, they trade off back from REM to AEM, most likely due to high coordination and expertise problems. The executive part of the board may evaluate the risk of accounting controversies to be considerably lower in these cases. **Figure 9** illustrates the course of the curve for reputation, low-CSR firms and the total sample firms, and it supports these considerations.

Figure 9 – The non-linear influence of BI on REMvsAEM

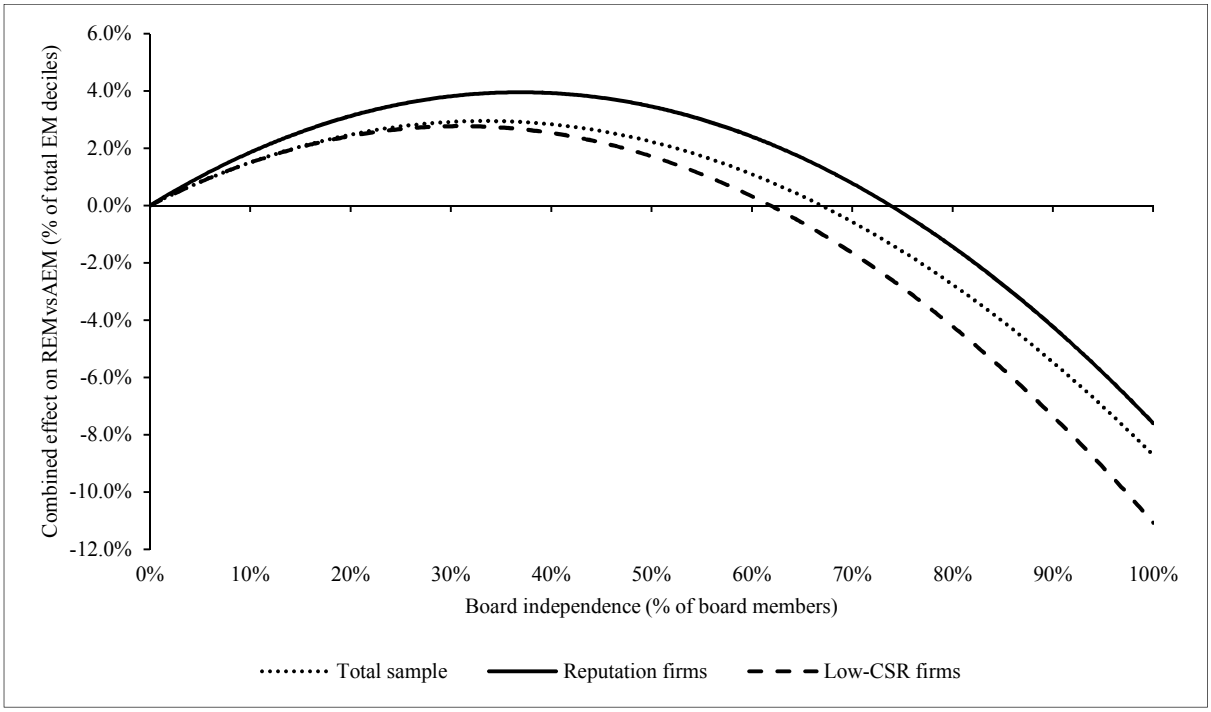


Figure illustrates the combined effect of the linear and squared coefficients for board independence on $\boldsymbol{REMvsAEM_{it}}$. The respective lines show the course for the total sample (2,733 obs.), high-reputation firms (1,033 obs.) and low-CSR firms (1,159 obs.) separately.

In summary, intrinsic CSR firms seem to further reduce their AEM actions as monitoring increases, and in particular, they refrain from using more REM instead of AEM, which can help them reach higher earnings without high risks of being detected, but may damage shareholder value. On the other hand, reputation enhancement firms seem to increase REM as AEM possibilities shrink, and therefore, both achieve their earnings goals and decrease the risk of investigations concerning their financial reports. This strategy comes at the possible expense of share- and stakeholders and affects future profitability. These findings indicate that CSR and, in particular, CSR incentives can moderate the effectiveness of the monitoring devices applied by shareholders. The respective CSR incentive influences the degree to which and whether firms trade off one type of EM for another that most likely has worse future effects for the user of the information than the type that was used in the first place. These results can be viewed as confirmation of $\boldsymbol{H_1}$, $\boldsymbol{H_2}$ and $\boldsymbol{H_3}$.

4.5 Additional analyses

4.5.1 Analysis of high-incentive situations

We follow the related literature and test whether our general results hold in situations that are thought to trigger higher incentives to pursue EM than the financial reporting decisions of normal years (e.g., Burgstahler & Dichev 1997; Graham et al. 2005; Roychowdhury 2006; Cohen et al. 2008). We test the zero earnings benchmark, which comprises disproportionate reactions from various stakeholders when the benchmark is not met. Similarly, future unfavourable negotiation power and equity valuation may cause the incentive to meet or beat the zero earnings. Second, analysts' forecasts can incite managers to take EM actions to provide the expected earnings numbers to capital market addressees (e.g., Roychowdhury 2006; Cohen et al. 2008; Bozzolan et al. 2015). Since these incentives can have higher importance than both reputational and intrinsic incentives, we check whether our results are distorted when we omit them. KPW and Chih et al. (2008) analyse the incentive situations on a univariate level and find a significantly lower use of EM for high-CSR firms. BFMM split their sample and re-estimate their regressions for all firms with above earnings forecast earnings.[163] We re-estimate **Equation 26** and **Equation 27**, adding $SUSPECT_{it}$, which is an indicator variable that captures the suspected earnings group or analyst forecast error interval.[164] **Table 45** provides the multivariate results, showing results that are qualitatively consistent with and quantitatively comparable to the findings stated above in **Table 43**.[165] This finding is not overall surprising since high-incentive situations should not bias the results of our questions about general EM strategies in different CSR and CG settings if our models are well specified. However, the zero earnings threshold seems to cause above average incentives that lead to higher AEM.

[163] However, firms with earnings that are considerably above the earnings forecasts may not have an incentive to manage their earnings. We alternatively apply an indicator variable when estimating the total sample to cancel out the possible omitted variable bias caused by the exclusion of high-incentive situations.
[164] We classify the full sample earnings and analyst forecast error intervals over the range from -0.075 to +0.075. The interval range is always 0.005. The suspect intervals are [0;0.005] for both situations. For the analyst forecast error, the general convention of the interval range of one cent is modified for a multi-currency setting. We additionally perform univariate tests following KPW. We find no significant differences in the EM proxies between high-CSR and low-CSR firms, indicating that all companies in the respective suspect situation use comparable amounts of EM.
[165] We additionally analyse the last year's earnings incentive and all three incentives for EM_ALL and $RAMvsAEM$ with and without CG characteristics. Untabulated results show that there is no change in our already stated results.

Table 45 – Multiple regressions of EM on CSR with SUSPECT control

	$SUSPECT_{it}$ *Zero Earnings*		$SUSPECT_{it}$ *Analyst Forecasts*	
	$\boldsymbol{DACC\_abs_{it}}$	$\boldsymbol{REM_{it}}$	$\boldsymbol{DACC\_abs_{it}}$	$\boldsymbol{REM_{it}}$
CSR_Score_{it}	-0.004	0.225	-0.004	0.224
	(3.89)***	(13.28)***	(3.82)***	(12.99)***
REP_{it}	0.023	0.045	0.023	0.045
	(6.27)***	(0.91)	(6.35)***	(0.90)
$CSR\_Score * REP_{it}$	-0.030	-0.093	-0.030	-0.092
	(7.70)***	(1.80)*	(7.75)***	(1.78)*
$SUSPECT_{it}$	0.003	0.011	0.000	0.043
	(1.99)**	(1.48)	(0.06)	(2.16)**
REM_{it}	0.003		0.003	
	(2.65)***		(2.64)***	
$DACC\_abs_{it}$		0.269		0.270
		(2.89)***		(2.87)***
$Size_{it-1}$	0.001	0.014	0.001	0.014
	(3.11)***	(5.95)***	(3.09)***	(5.86)***
MB_{it-1}	0.000	0.013	0.000	0.013
	(0.48)	(5.18)***	(0.45)	(5.30)***
$LOSS_{it}$	0.020	0.024	0.020	0.024
	(23.99)***	(3.67)***	(24.56)***	(3.36)***
$Leverage_{it-1}$	-0.012	-0.096	-0.012	-0.097
	(6.19)***	(6.59)***	(6.19)***	(6.84)***
$\Delta EMPLOY_{it}$	0.001	-0.049	0.001	-0.049
	(1.20)	(2.67)***	(1.14)	(2.75)***
ΔTA_{it}	0.000	-0.022	0.000	-0.022
	(0.69)	(1.00)	(0.76)	(1.01)
GOV_Score_{it}	0.006	-0.055	0.006	-0.056
	(3.32)***	(3.78)***	(3.19)***	(3.84)***
ACC_{it-1}	-0.016	-0.258	-0.016	-0.264
	(6.19)***	(8.64)***	(6.34)***	(8.66)***
CFO_{it}	0.001	1.282	0.000	1.281
	(0.15)	(40.00)***	(0.05)	(39.28)***
ATO_{it-1}	0.000	-0.036	0.000	-0.036
	(0.80)	(12.80)***	(0.77)***	(12.65)***
$Litigation_{it}$	-0.000	0.157	-0.000	0.157
	(0.19)	(16.64)***	(0.19)***	(16.69)***
ΔGDP_{jt}	-0.000	-0.044	0.000	-0.047
	(0.03)	(0.81)	(0.02)	(0.91)
$FirmAge_{it}$	-0.001	0.031	-0.001	0.031
	(2.36)**	(5.85)***	(2.34)**	(5.81)***
$Risk_{it}$	0.002	0.028	0.002	0.028
	(4.44)***	(6.26)***	(4.53)***	(6.27)***
Year Dummies	Included	Included	Included	Included
Industry Dummies	Included	Included	Included	Included
R^2	0.23	0.26	0.23	0.26
N	2,733	2,733	2,733	2,733

*, **, and *** indicate significance at the 0.10, 0.05, and 0.01 levels, respectively, based on two-tailed tests. All test statistics are calculated using Driscoll-Kraay standard errors with lag (7). The mean VIF for all regressions are again considerably below 5. $SUSPECT_{it}$ is an indicator for the respective high-incentive situation, for the zero earnings threshold with the earnings interval [0;0.005] and for the forecasted earnings threshold with the analyst forecast error interval [0;0.005]. See Appendix D for detailed variable descriptions.

4.5.2 Different discretionary accruals and real activities EM measures

To further test the robustness of our results, we conduct all analyses with alternative discretionary accruals from the performance-matched modified Jones model (DeFond & Subramanyam 1998; Kothari et al. 2005) and the performance-matched adapted Jones model (Dechow et al. 2003). In addition, following Cohen and Zarowin (2010) and Zang (2012), we apply two alternative combined proxies for REM to separate the ambiguous effects of abnormal operating cash flows. First, we comprehend REM_{it} as the sum of abnormal discretionary expenses and abnormal production costs, multiplied by negative one.[166] Alternatively, we obtain REM_{it} as the sum of abnormal operating cash flows and abnormal discretionary expenses and use only positive values of REM_{it} due to the predominant income-increasing nature of the discussed situations. We proceed equivalently for a modified calculation of $REMvsAEM_{it}$ and EM_ALL_{it}. The results remain consistent with the tabulated results. Alternatively, we follow KPW and use separate proxies for positive and negative discretionary accruals (AEM) and separate proxies for all types of REM in our regressions. Further, we use standardized values of all EM proxies (e.g., Cohen et al. 2008). The results are similarly qualitatively equal, quantitatively even better in some cases, and therefore further support our findings.[167]

4.5.3 Different CSR firm criteria and additional control variables

Concerning the definition of CSR incentives, analysis of board characteristics and suspect intervals on EM, we alternatively define high-CSR firms as firms in the top third and low-CSR firms as firms in the bottom third of the CSR distribution. This leads to a more conservative distinction between high- and low-CSR firms. The results remain qualitatively the same for all regressions. Since CSR incentives may exist only when CSR investments exceed a certain amount, we also alter the general formulation and use only indicators for firms with above median or top third CSR performance. Our results remain, although there is a loss in efficiency, because there is no longer any continuous variable for CSR. We incorporate additional control variables to cancel out an omitted variable bias. The integration of a dummy for financial statements audited by Big 4 companies

[166] The negative scaling remains for the combined real activities EM proxies for reasons of consistency.
[167] For reasons of manageable size, the regression tables for these robustness checks remain untabulated.

does not alter our results, nor does an additional control for the board size and/or number of board meetings. Further, we use a performance control to rule out the possibility of biased coefficients due to a solely implicit performance control by the applied variables and models. The inclusion of return on assets and industry-mean adjusted return on assets as a separate regressor leaves the results qualitatively the same and quantitatively almost equal to the already tabulated findings. These observations confirm our specification and proposition of a sufficient performance control immediately from the outset.[168]

4.6 Conclusion

This study analyses the influence of CSR on EM and the effect that CG has on EM. In particular, the study tries to elucidate the diverse incentives for CSR investments. We therefore ask whether firms that invest in CSR engagement also convey doing so as a cooperative strategy in their EM motivations or whether they use it as a vehicle to boost their image and conformably decide about EM. Furthermore, we question whether these diverse incentives have a moderating role in whether independence in the board fosters monitoring and therefore reporting transparency and shareholder value.

The outcomes show consistent proof for diverse relationships between CSR performance and EM, depending on the predominant incentive behind CSR investment. For the predominantly intrinsic segment of CSR firms, higher CSR entails a thorough approach with lower types of EM and even a shift from REM to AEM to prevent shareholder value from being damaged. For the predominantly reputational segment of CSR firms, the primary goal is an improvement of their image, which is why they decide to use lower AEM to ensure that they preclude investigations and reputational damage, while they use even more REM and thereby damage future shareholder value. We further find that higher percentages of independent board members upgrade monitoring processes to a certain degree, improving the degree of transparent reporting in terms of lower AEM. However, firms that predominantly seek to improve their reputation take the easy route and use even

[168] The results for all these alterations again remain untabulated.

more REM, which the board can hardly detect or even prevent, and thereby jeopardize shareholder value.

Our findings hold after controlling for various control variables in the multivariate analyses, the integration of a special situations control and an additional control for board specifics, performance and auditing. They are robust to various types of discretionary accruals and real activities measures, the separate use of proxies as well as EM proxy standardization. Further, they are robust to an alternative separation of the sample into firms of high and low CSR performance and sole use of indicators in formulation of the CSR incentives. Overall, the results are consistent with $\boldsymbol{H_1}$ and $\boldsymbol{H_2}$. The findings on board independence confirm $\boldsymbol{H_3}$.

Naturally, the study has some limitations. In particular, there may be more than the already controlled high-incentive situations, which could influence the results. Additionally, possible influencing effects from the levels of enforcement in the different countries are captured only by the industry and year instruments and other control variables and were not specified separately.[169] This could lead to changes in the conclusions. The general distinction between one-tier and two-tier systems concerning board structure can slightly affect the comparability of results between the distinct systems due to diverse monitoring mechanisms and opportunities to influence financial reporting.[170] The general criticism of EM modelling using residuals of first-stage regressions, particularly for discretionary accruals, also holds for this study. However, we try to minimize the frequently discussed measurement errors by choosing the most sophisticated existent models, estimating them cross-sectionally and controlling for performance. Similarly, differentiation of the two incentive groups using the CSR reporting behaviour may not lead to a perfect distinction, and therefore, the results may lack efficiency. Nevertheless, the results of our study seem to prove that voluntary CSR disclosure can be a proper vehicle for differentiation. Subsequent studies can focus on disentangling the mediating role of CSR incentives concerning other board characteristics, the auditing process or regulation. We may therefore gain further insights into what features drive EM as well as the trade-off relationship

[169] On the other hand, we try to minimize the possibility of omitted variable bias to a very high degree by including the various controls.
[170] However, we regard these differences as no significant influence on the results of the study (e.g., Jungmann 2007). In particular, the distinction may not overall ground the possibilities to pursue independence in collocation.

and how CSR incentives correspond. This could be of interest for the addressees of disclosures, lead to a better understanding of which instruments support transparency and enable stakeholders to make useful decisions. Additionally, it may be interesting to evaluate future performance following high CSR performance for both CSR incentives, particularly in regard to whether high-reputation firms succeed in achieving competitive advantages and higher future economic performance.

5. References

Abbott, L.J., Parker, S., Peters, G.F., 2006. Earnings Management, Litigation Risk, and Asymmetric Audit Fee Responses. AUDITING: A Journal of Practice & Theory 25 (1), 85–98. 10.2308/aud.2006.25.1.85.

Abbott, L.J., Daugherty, B., Parker, S., Peters, G.F., 2016. Internal Audit Quality and Financial Reporting Quality: The Joint Importance of Independence and Competence. Journal of Accounting Research 54 (1), 3–40. 10.1111/1475-679X.12099.

Acharya, V.V., Ryan, S.G., 2016. Banks' Financial Reporting and Financial System Stability. Journal of Accounting Research 54 (2), 277–340. 10.1111/1475-679X.12114.

Agarwal, S., Chomsisengphet, S., Liu, C., Ghon Rhee, S., 2007. Earnings management behaviors under different economic environments: Evidence from Japanese banks. International Review of Economics & Finance 16 (3), 429–443. 10.1016/j.iref.2005.08.003.

Ahmed, A.S., 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. 10.1016/S0165-4101(99)00017-8.

Ali, A., Zhang, W., 2015. CEO tenure and earnings management. Journal of Accounting and Economics 59 (1), 60–79. 10.1016/j.jacceco.2014.11.004.

Anandarajan, A., Hasan, I., McCarthy, C., 2007. Use of loan loss provisions for capital, earnings management and signalling by Australian banks. Accounting & Finance 47 (3), 357–379. 10.1111/j.1467-629x.2007.00220.x.

Anderson, M., Banker, R., Janakiraman, S., 2003. Are Selling, General, and Administrative Costs "Sticky"? Journal of Accounting Research 41 (1), 47–63.

Antle, R., Gordon, E., Narayanamoorthy, G., Zhou, L., 2006. The joint determination of audit fees, non-audit fees, and abnormal accruals. Rev Quant Finan Acc 27 (3), 235–266. 10.1007/s11156-006-9430-y.

Arellano, M., 1987. Practitioners' Corner: Computing Robust Standard Errors for Within-groups Estimators*. Oxford Bulletin of Economics and Statistics 49 (4), 431–434. 10.1111/j.1468-0084.1987.mp49004006.x.

Arellano, M., Bover, O., 1995. Another look at the instrumental variable estimation of error-components models. Journal of Econometrics 68 (1), 29–51. 10.1016/0304-4076(94)01642-D.

Armstrong, C., Guay, W., Weber, J., 2010. The role of information and financial reporting in corporate governance and debt contracting. Journal of Accounting and Economics 50, 179–234. 10.1016/j.jacceco.2010.10.001.

Ashbaugh, H., LaFond, R., Mayhew, B.W., 2003. Do Nonaudit Services Compromise Auditor Independence?: Further Evidence. The Accounting Review 78 (3), 611–639. 10.2308/accr.2003.78.3.611.

Atkins, B., 2006. Is corporate social responsibility responsible? Forbes.com. Available at https://www.forbes.com/2006/11/16/leadership-philanthropy-charity-lead-citizen-cx_ba_1128directorship.html. Updated on November 28 2006. Checked on November 27 2017.

Badertscher, B.A., 2011. Overvaluation and the Choice of Alternative Earnings Management Mechanisms. The Accounting Review 86 (5), 1491–1518. 10.2308/accr-10092.

Badolato, P.G., Donelson, D.C., Ege, M., 2014. Audit committee financial expertise and earnings management: The role of status. Journal of Accounting and Economics 58 (2-3), 208–230. 10.1016/j.jacceco.2014.08.006.

Ball, R., Shivakumar, L., 2005. Earnings quality in UK private firms: Comparative loss recognition timeliness. Journal of Accounting and Economics 39 (1), 83–128. 10.1016/j.jacceco.2004.04.001.

Ball, R., Shivakumar, L., 2006. The Role of Accruals in Asymmetrically Timely Gain and Loss Recognition. Journal of Accounting Research 44 (2), 207–242. 10.1111/j.1475-679X.2006.00198.x.

Ballwieser, W., Grewe, W. (Eds.), 2008. Wirtschaftsprüfung im Wandel - Herausforderungen an Wirtschaftsprüfung, Steuerberatung, Consulting und Corporate Finance, 1st ed. Beck Publishing, München.

Bauer, M., 2005. Die Unabhängigkeit des Abschlussprüfers im Zusammenhang mit dem gleichzeitigen Angebot von Beratungsleistungen beim Prüfungsmandanten: Eine empirische Analyse. Inauguraldissertation. Opus Bayern, Wuerzburg.

Beasley, M.S., 1996. An Empirical Analysis of the Relation between the Board of Director Composition and Financial Statement Fraud. The Accounting Review 71 (4), 443–465.

Beatty, A., Chamberlain, S.L., Magliolo, J., 1995. Managing Financial Reports of Commercial Banks: The Influence of Taxes, Regulatory Capital, and Earnings. Journal of Accounting Research 33 (2), 231. 10.2307/2491487.

Beatty, A., Ke, B., Petroni, K., 2002. Earnings Management to Avoid Earnings Declines across Publicly and Privately Held Banks. The Accounting Review 77 (3), 547–570.

Beatty, A., Liao, S., 2011. Do delays in expected loss recognition affect banks' willingness to lend? Journal of Accounting and Economics 52 (1), 1–20. 10.1016/j.jacceco.2011.02.002.

Beatty, A., Liao, S., 2014. Financial accounting in the banking industry: A review of the empirical literature. Journal of Accounting and Economics 58 (2-3), 339–383. 10.1016/j.jacceco.2014.08.009.

Beatty, A.L., Harris, D.G., 1998. The Effects of Taxes, Agency Costs and Information Asymmetry on Earnings Management: A Comparison of Public and Private Firms. Review of Accounting Studies 3, 299–326.

Beatty, A.L., Liao, S., 2009. Regulatory Capital Ratios, Loan Loss Provisioning and Pro-Cyclicality. SSRN Journal. 10.2139/ssrn.1463374.

Beaver, W.H., Engel, E.E., 1996. Discretionary behavior with respect to allowances for loan losses and the behavior of security prices. Journal of Accounting and Economics 22 (1-3), 177–206. 10.1016/S0165-4101(96)00428-4.

Beck, P.J., Narayanamoorthy, G.S., 2013. Did the SEC impact banks' loan loss reserve policies and their informativeness? Journal of Accounting and Economics 56 (2-3), 42–65. 10.1016/j.jacceco.2013.06.002.

Becker, C.L., DeFond, M.L., Jiambalvo, J., Subramanyam, K.R., 1998. The Effect of Audit Quality on Earnings Management. Contemporary Accounting Research 15 (1), 1–24. 10.1111/j.1911-3846.1998.tb00547.x.

Bergstresser, D., Philippon, T., 2006. CEO incentives and earnings management. Journal of Financial Economics 80 (3), 511–529. 10.1016/j.jfineco.2004.10.011.

Bigus, J., Zimmermann, R.-C., 2009. Quasirentenmodell und Honorare für Abschlussprüfungen in Deutschland – eine empirische Analyse. Z Betriebswirtsch 79 (11), 1283–1308. 10.1007/s11573-009-0317-y.

Bikker, J.A., Metzemakers, P.A.J., 2005. Bank provisioning behaviour and procyclicality. Journal of International Financial Markets, Institutions and Money 15 (2), 141–157. 10.1016/j.intfin.2004.03.004.

Bischof, S., 2006. Anhangangaben zu den Honoraren für Leistungen des Abschlussprüfers. Die Wirtschaftsprüfung 59 (11), 705–713.

Blundell, R., Bond, S., 1998. Initial conditions and moment restrictions in dynamic panel data models. Journal of Econometrics 87 (1), 115–143. 10.1016/S0304-4076(98)00009-8.

Bonner, S.E., 2008. Judgment and Decision Making in Accounting. Englewood Cliffs, New Jersey: Prentice Hall.

Botosan, C.A., 1997. Disclosure level and the cost of equity capital. The Accounting Review 72 (3), 323–349.

Botosan, C.A., Plumlee, M.A., 2002. A re-examination of disclosure level and the expected cost of equity capital. Journal of Accounting Research 40 (1), 21–40. 10.1111/1475-679X.00037.

Bouvatier, V., Lepetit, L., Strobel, F., 2014. Bank income smoothing, ownership concentration and the regulatory environment. Journal of Banking & Finance 41, 253–270. 10.1016/j.jbankfin.2013.12.001.

Bowen, R.M., DuCharme, L., Shores, D., 1995. Stakeholders' implicit claims and accounting method choice. Journal of Accounting and Economics 20 (3), 255–295. 10.1016/0165-4101(95)00404-1.

Bozzolan, S., Fabrizi, M., Mallin, C.A., Michelon, G., 2015. Corporate Social Responsibility and Earnings Quality: International Evidence. The International Journal of Accounting 50 (4), 361–396. 10.1016/j.intacc.2015.10.003.

Bradbury, M., Mak, Y.T., Tan, S.M., 2006. Board Characteristics, Audit Committee Characteristics and Abnormal Accruals. Pacific Accounting Review 18 (2), 47–68. 10.1108/01140580610732813.

Bratten, B., Jennings, R., Schwab, C.M., 2016. The accuracy of disclosures for complex estimates: Evidence from reported stock option fair values. Accounting, Organizations and Society 52, 32–49. 10.1016/j.aos.2015.09.001.

Breusch, T.S., Pagan, A.R., 1980. The Lagrange Multiplier Test and its Applications to Model Specification in Econometrics. The Review of Economic Studies 47 (1), 239. 10.2307/2297111.

Burgstahler, D., Dichev, I., 1997. Earnings management to avoid earnings decreases and losses. Journal of Accounting and Economics 24 (1), 99–126. 10.1016/S0165-4101(97)00017-7.

Bushman, R.M., 2014. Thoughts on financial accounting and the banking industry. Journal of Accounting and Economics 58 (2-3), 384–395. 10.1016/j.jacceco.2014.09.004.

Bushman, R.M., Williams, C.D., 2012. Accounting discretion, loan loss provisioning, and discipline of Banks' risk-taking. Journal of Accounting and Economics 54 (1), 1–18. 10.1016/j.jacceco.2012.04.002.

Bushman, R.M., Williams, C.D., 2015. Delayed Expected Loss Recognition and the Risk Profile of Banks. Journal of Accounting Research 53 (3), 511–553. 10.1111/1475-679X.12079.

Cahan, S., Emanuel, D., Hay, D., Wong, N., 2008. Non-audit fees, long-term auditor–client relationships and earnings management. Accounting & Finance 48 (2), 181–207. 10.1111/j.1467-629X.2008.00251.x.

Calabrese, A., Costa, R., Levialdi, N., Menichini, T., 2016. A fuzzy analytic hierarchy process method to support materiality assessment in sustainability reporting. Journal of Cleaner Production 121, 248–264. 10.1016/j.jclepro.2015.12.005.

Calegari, M.F., Chotigeat, T., Harjoto, M., 2010. Corporate social responsibility and earnings reporting. Journal of Current Research in Global Business 13 (20), 1–14.

Callan, S.J., Thomas, J.M., 2009. Corporate financial performance and corporate social performance: An update and reinvestigation. Corp. Soc. Responsib. Environ. Mgmt 16 (2), 61–78. 10.1002/csr.182.

Cameron, A.C., Gelbach, J.B., Miller, D.L., 2011. Robust Inference With Multiway Clustering. Journal of Business & Economic Statistics 29 (2), 238–249.

Cameron, A.C., Trivedi, P.K., 2010. Microeconometrics using Stata. Texas: Stata Press.

Carroll, A.B., 1979. A Three-Dimensional Conceptual Model of Corporate Performance. Academy of Management Review 4 (4), 497–505. 10.5465/amr.1979.4498296.

Cavallo, M., Majnoni, G., 2002. Do Banks Provision for Bad Loans in Good Times? Empirical Evidence and Policy Implications, in: Levich, R.M., Majnoni, G., Reinhart, C.M. (Eds.), Ratings, Rating Agencies and the Global Financial System, vol. 9. Springer US, Boston, MA, pp. 319–342.

CFA Institute, 2013. Financial Reporting Disclosures – Investor Perspectives on Transparency, Trust, and Volume. Available at https://www.cfapubs.org/doi/pdf/10.2469/ccb.v2013.n12.1. Updated on July 2013. Checked on November 29 2017.

Chatterji, A.K., Durand, R., Levine, D.I., Touboul, S., 2016. Do ratings of firms converge?: Implications for managers, investors and strategy researchers. Strat. Mgmt. J. 37 (8), 1597–1614. 10.1002/smj.2407.

Chen, T.-K., Tseng, Y., Hsieh, Y.-T., 2014. Real Earnings Management Uncertainty and Corporate Credit Risk. European Accounting Review 24 (3), 413–440. 10.1080/09638180.2014.918518.

Cheng, B., Ioannou, I., Serafeim, G., 2014. Corporate social responsibility and access to finance. Strat. Mgmt. J. 35 (1), 1–23. 10.1002/smj.2131.

Cheng, Q., Warfield, T.D., 2005. Equity Incentives and Earnings Management. The Accounting Review 80 (2), 441–476. 10.2308/accr.2005.80.2.441.

Cheng, Q., Warfield, T.D., Ye, M., 2011. Equity Incentives and Earnings Management: Evidence from the Banking Industry. Journal of Accounting, Auditing & Finance 26, 317-349.

Chi, W., Douthett, E.B., Lisic, L.L., 2012. Client importance and audit partner independence. Journal of Accounting and Public Policy 31 (3), 320–336. 10.1016/j.jaccpubpol.2011.08.009.

Chi, W., Lisic, L.L., Pevzner, M., 2011. Is Enhanced Audit Quality Associated with Greater Real Earnings Management? Accounting Horizons 25 (2), 315–335. 10.2308/acch-10025.

Chih, H.-L., Shen, C.-H., Kang, F.-C., 2008. Corporate Social Responsibility, Investor Protection, and Earnings Management: Some International Evidence. J Bus Ethics 79 (1-2), 179–198. 10.1007/s10551-007-9383-7.

Cho, E., Chun, S., 2016. Corporate social responsibility, real activities earnings management, and corporate governance: Evidence from Korea. Asia-Pacific Journal of Accounting & Economics 23 (4), 400–431. 10.1080/16081625.2015.1047005.

Choi, B.B., Lee, D., Park, Y., 2013. Corporate Social Responsibility, Corporate Governance and Earnings Quality: Evidence from Korea. Corporate Governance: An International Review 21 (5), 447–467. 10.1111/corg.12033.

Chun, S.B., Kwon, H.J., Kim, S.H., Lee, A.Y., 2012. Financial reporting patterns of korean firms: Based on the survey of chief financial officers. Korean Accounting Journal 21 (6), 97–132.

Chung, H., Kallapur, S., 2003. Client Importance, Nonaudit Services, and Abnormal Accruals. The Accounting Review 78 (4), 931–955.

Clor-Proell, S.M., Proell, C.A., Warfield, T.D., 2014. The effects of presentation salience and measurement subjectivity on nonprofessional investors' fair value judgments. Contemporary Accounting Research 31 (1), 45–66. 10.1111/1911-3846.12041.

Cohen, L.J.E.E., Cornett, M.M., Marcus, A.J., Tehranian, H., 2014. Bank Earnings Management and Tail Risk during the Financial Crisis. Journal of Money, Credit and Banking 46 (1), 171–197. 10.1111/jmcb.12101.

Cohen, D.A., Dey, A., Lys, T.Z., 2008. Real and Accrual-Based Earnings Management in the Pre- and Post-Sarbanes-Oxley Periods. The Accounting Review 83 (3), 757–787. 10.2308/accr.2008.83.3.757.

Cohen, D.A., Zarowin, P., 2010. Accrual-based and real earnings management activities around seasoned equity offerings. Journal of Accounting and Economics 50 (1), 2–19. 10.1016/j.jacceco.2010.01.002.

Cohen, W.A., 2009. What Drucker taught us about social responsibility. Leader to Leader 2009 (51), 29–34. 10.1002/ltl.315.

Collins, J.H., Shackelford, D.A., Wahlen, J.M., 1995. Bank Differences in the Coordination of Regulatory Capital, Earnings, and Taxes. Journal of Accounting Research 33 (2), 263. 10.2307/2491488.

Commerford, B.P., Hermanson, D.R., Houston, R.W., Peters, M.F., 2016. Real Earnings Management: A Threat to Auditor Comfort? AUDITING: A Journal of Practice & Theory 35 (4), 39–56. 10.2308/ajpt-51405.

Coulton, J.J., Ruddock, C., 2012. Client importance and audit reporting quality at the partner, office and firm level. Programme & Proceedings 35th Annual Congress of the European Accounting Association.

Craswell, A., Stokes, D.J., Laughton, J., 2002. Auditor independence and fee dependence. Journal of Accounting and Economics 33 (2), 253–275. 10.1016/S0165-4101(02)00044-7.

Dalton, D.R., Daily, C.M., Johnson, J.J., Ellstrand, A.E., 1999. Number of Directors and Financial Performance: A Meta-Analysis. The Academy of Management Journal 42 (6), 674–686.

Dalton, D.R., Dalton, C.M., 2010. Women and corporate boards of directors: The promise of increased, and substantive, participation in the post Sarbanes-Oxley era. Business Horizons 53 (3), 257–268. 10.1016/j.bushor.2009.12.004.

Davidson, R., Goodwin-Stewart, J., Kent, P., 2005. Internal governance structures and earnings management. Accounting & Finance 45 (2), 241–267. 10.1111/j.1467-629x.2004.00132.x.

Davidson, R., and MacKinnon, J.G. (1993). Estimation and inference in econometrics. New York: Oxford University Press.

Davidson, W.N., Jiraporn, P., Nemec, C., 2004. Earnings Management Following Duality-Creating Successions: Ethnostatistics, Impression Management, And Agency Theory. Academy of Management Journal 47 (2), 267–275. 10.2307/20159577.

DeAngelo, L.E., 1981a. Auditor independence, 'low balling', and disclosure regulation. Journal of Accounting and Economics 3 (2), 113–127. 10.1016/0165-4101(81)90009-4.

DeAngelo, L.E., 1981b. Auditor size and audit quality. Journal of Accounting and Economics 3 (3), 183–199. 10.1016/0165-4101(81)90002-1.

DeAngelo, L., 1986. Accounting numbers as market valuation substitutes: A study of management buyouts of public stockholders. The Accounting Review 61 (3), 400–420.

DeBoskey, D.G., Jiang, W., 2012. Earnings management and auditor specialization in the post-sox era: An examination of the banking industry. Journal of Banking & Finance 36 (2), 613–623. 10.1016/j.jbankfin.2011.09.007.

Dechow, P.M., 1994. Accounting earnings and cash flows as measure of firm performance – The role of accounting accruals. Journal of Accounting and Economics 18 (1), 3–42. 10.1016/0165-4101(94)90016-7.

Dechow, P.M., Ge, W., Schrand, C., 2010. Understanding earnings quality: A review of the proxies, their determinants and their consequences. Journal of Accounting and Economics 50 (2-3), 344–401. 10.1016/j.jacceco.2010.09.001.

Dechow, P.M., Dichev, I.D., 2002. The Quality of Accruals and Earnings: The Role of Accrual Estimation Errors. The Accounting Review 77 (s-1), 35–59. 10.2308/accr.2002.77.s-1.35.

Dechow, P.M., Kothari, S.P., L. Watts, R., 1998. The relation between earnings and cash flows. Journal of Accounting and Economics 25 (2), 133–168. 10.1016/s0165-4101(98)00020-2.

Dechow, P.M., Richardson, S.A., Tuna, I., 2003. Why are earnings kinky? An examination of the earnings management explanation. Review of Accounting Studies 8 (2/3), 355–384. 10.1023/a:1024481916719.

Dechow, P.M., Sloan, R.G., 1991. Executive incentives and the horizon problem. Journal of Accounting and Economics 14 (1), 51–89. 10.1016/0167-7187(91)90058-S.

Dechow, P.M., Sloan, R.G., Sweeny, A.P., 1995. Detecting Earnings Management. The Accounting Review 70 (2), 193–225.

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. 10.1111/j.1911-3846.1996.tb00489.x.

DeFond, M.L., Francis, J.R., Wong, T.J., 2000. Auditor Industry Specialization and Market Segmentation: Evidence from Hong Kong. AUDITING: A Journal of Practice & Theory 19 (1), 49–66. 10.2308/aud.2000.19.1.49.

DeFond, M.L., Jiambalvo, J., 1993. Factors Related to Auditor-Client Disagreements over Income-Increasing Accounting Methods. Contemporary Accounting Research 9 (2), 415–431. 10.1111/j.1911-3846.1993.tb00889.x.

DeFond, M.L., Raghunandan, K., Subramanyam, K.R., 2002. Do Non-Audit Service Fees Impair Auditor Independence?: Evidence from Going Concern Audit Opinions. Journal of Accounting Research 40 (4), 1247–1274. 10.1111/1475-679X.00088.

DeFond, M.L., Subramanyam, K.R., 1998. Auditor changes and discretionary accruals. Journal of Accounting and Economics 25 (1), 35–67. 10.1016/s0165-4101(98)00018-4.

DeFond, M.L., Zhang, J., 2014. A review of archival auditing research. Journal of Accounting and Economics 58 (2-3), 275–326. 10.1016/j.jacceco.2014.09.002.

DeHoyos, R.E., Sarafidis, V., 2006. Testing for cross-sectional dependence in panel-data models. The Stata Journal 6 (4), 482–496.

DeMaCarty, P., 2009. Financial Returns of Corporate Social Responsibility, and the Moral Freedom and Responsibility of Business Leaders. Business and Society Review 114 (3), 393–433. 10.1111/j.1467-8594.2009.00348.x.

Deutscher Bundestag, 2004a. BT-Drucksache 15/3419: Entwurf eines Gesetzes zur Einführung internationaler Rechnungslegungsstandards und zur Sicherung der Qualität der Abschlussprüfung: Bilanzrechtsreformgesetz - BilReG.

Deutscher Bundestag, 2004b. BT-Drucksache 15/4054: Entwurf eines Gesetzes zur Einführung internationaler Rechnungslegungsstandards und zur Sicherung der Qualität der Abschlussprüfung: Bilanzrechtsreformgesetz - BilReG.

Deutscher Bundestag, 2006. BT-Drucksache 16/2858: Entwurf eines Gesetzes zur Stärkung der Berufsaufsicht und zur Reform berufsrechtli-cher Regelungen in der Wirtschaftsprüferordnung: Berufsaufsichtsreformgesetz - BARefG.

Deutscher Bundestag, 2007. BT-Drucksache 16/10067: Entwurf eines Gesetzes zur Modernisierung des Bilanzrechts: Bilanzrechtsmodernisierungsgesetz - BilMoG.

Deutscher Bundestag, 2007. Gesetz zur Stärkung der Berufsaufsicht und zur Reform berufsrechtlicher Regelungen in der Wirtschaftsprüferordnung: Berufsaufsichtsreformgesetz - BARefG.

Deutscher Bundestag, 2009. BT-Drucksache 16/12407: Entwurf eines Gesetzes zur Modernisierung des Bilanzrechts: Bilanzrechtsmodernisierungsgesetz - BilMoG.

Diamond, D., 1984 Financial intermediation and delegated monitoring. The Review of Economic Studies 51, 393–414. 10.2307/2297430.

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. 10.1016/j.jacceco.2013.05.004.

Driscoll, J.C., Kraay, A.C., 1998. Consistent Covariance Matrix Estimation with Spatially Dependent Panel Data. Review of Economics and Statistics 80 (4), 549–560. 10.1162/003465398557825.

Ebke, W.F., 2008. Kommentierung zu §§ 319, 319a, in: Schmidt, K., Ebke, W.F. (Eds.), Münchener Kommentar zum Handelsgesetzbuch, 2nd ed. Beck Publishing, München.

El Sood, H.A., 2012. Loan loss provisioning and income smoothing in US banks pre and post the financial crisis. International Review of Financial Analysis 25, 64–72. 10.1016/j.irfa.2012.06.007.

European Commission, 1996. Green Paper - The Role, the Position and the Liability of the Statutory Auditor within the European Union. European Commission, Brussels.

European Commission, 2010. Green Paper - Audit Policy: Lessons from the Crisis. European Commission, Brussels.

European Union, 2006. Directive 2006/43/EC of the European Parliament and of the Council on statutory audits of annual accounts and consolidated accounts, amending Council Directives 78/660/EEC and 83/349/EEC and repealing Council Directive 84/253/EEC.

European Union, 2014. Regulation (EU) No 537/2014 of the European Parliament and of the Council on specific requirements regarding statutory audit of public-interest entities and repealing Commission Decision 2005/909/EC: L157, 77-112.

Evans, M.E., Houston, R.W., Peters, M.F., Pratt, J.H., 2015. Reporting Regulatory Environments and Earnings Management: U.S. and Non-U.S. Firms Using U.S. GAAP or IFRS. The Accounting Review 90 (5), 1969–1994. 10.2308/accr-51008.

Ewert, R., 2004. Audit Regulation, Audit Quality, and Audit Research in the Post-Enron Era: An Analysis of Nonaudit Services, in: Leuz, C., Pfaff, D., Hopwood, A. (Eds.), The Economics and Politics of Accounting. Oxford University Press, Oxford, 239-.

Fama, E.F., Jensen, M.C., 1983. Separation of Ownership and Control. The Journal of Law and Economics 26 (2), 301–325. 10.1086/467037.

Farber, D.B., 2005. Restoring Trust after Fraud: Does Corporate Governance Matter? The Accounting Review 80 (2), 539–561. 10.2308/accr.2005.80.2.539.

Federal Financial Institutions Examination Council (FFIEC) (2001). Policy Statement on Allowance for Loan and Lease Losses (ALLL) Methodologies and Documentation for Banks and Savings Institutions. July 6.

Ferguson, M.J., Seow, G.S.I.M., Young, D., 2004. Nonaudit Services and Earnings Management: UK Evidence. Contemporary Accounting Research 21 (4), 813–841. 10.1506/MFV5-9T3Q-H5RK-VC20.

Fieller, E.C., 1954. Some problems in interval estimation. Journal of the Royal Statistical Society. Series B (Methodological) 16 (2), 175–185.

Financial Reporting Council (FRC), 2016. Revised Ethical Standard 2016: Ethical Standard (2016).

Financial Stability Forum (2009).Report of the financial stability forum on addressing procyclicality in the financial system. Available at http://www.financialstabilityboard.org/publications/r_0904a.pdf. Updated on April 2 2009. Checked on November 29 2017.

Fonseca, A.R., González, F., 2008. Cross-country determinants of bank income smoothing by managing loan-loss provisions. Journal of Banking & Finance 32 (2), 217–228. 10.1016/j.jbankfin.2007.02.012.

Francis, J.R., 2004. What do we know about audit quality? The British Accounting Review 36 (4), 345–368. 10.1016/j.bar.2004.09.003.

Francis, J.R., Maydew, E.L., Sparks, H.C., 1999. The Role of Big 6 Auditors in the Credible Reporting of Accruals. AUDITING: A Journal of Practice & Theory 18 (2), 17–34. 10.2308/aud.1999.18.2.17.

Frankel, R.M., Johnson, M.F., Nelson, K.K., 2002. The Relation between Auditors' Fees for Nonaudit Services and Earnings Management. The Accounting Review 77 (s-1), 71–105. 10.2308/accr.2002.77.s-1.71.

Freixas, X., Rochet, J.-C., 2008. Microeconomics of Banking, 2nd ed. Cambridge, Massachusetts: MIT Press.

Froot, K.A., 1989. Consistent Covariance Matrix Estimation with Cross-Sectional Dependence and Heteroskedasticity in Financial Data. The Journal of Financial and Quantitative Analysis 24 (3), 333. 10.2307/2330815.

Garriga, E., Melé, D., 2004. Corporate Social Responsibility Theories: Mapping the Territory. J Bus Ethics 53 (1/2), 51–71. 10.1023/b:busi.0000039399.90587.34.

Gastinger, K., Gaggl, P., 2015. CSR und Berichterstattung. in: Schneider, A., Schmidtpeter, R. (Eds.), Corporate Social Responsibility – Verantwortungsvolle Unternehmensführung in Theorie und Praxis. 2nd ed. Springer Gabler Publishing, Berlin Heidelberg.

Gaynor, L.M., Kelton, A.S., Mercer, M., Yohn, T.L., 2016. Understanding the Relation between Financial Reporting Quality and Audit Quality. AUDITING: A Journal of Practice & Theory 35 (4), 1–22. 10.2308/ajpt-51453.

Gebhardt, G., Novotny-Farkas, Z., 2011. Mandatory IFRS Adoption and Accounting Quality of European Banks. Journal of Business Finance & Accounting 38 (3-4), 289–333. 10.1111/j.1468-5957.2011.02242.x.

Geiger, M.A., Rama, D.V., 2003. Audit Fees, Nonaudit Fees, and Auditor Reporting on Stressed Companies. AUDITING: A Journal of Practice & Theory 22 (2), 53–69. 10.2308/aud.2003.22.2.53.

Global Reporting Initiative (GRI), 2017. Getting Started with the GRI standards. Available at https://www.globalreporting.org/standards/getting-started-with-the-gri-standards/. Checked on November 27 2017.

Gow, I.D., Larcker, D.F., Reiss, P.C., 2016. Causal Inference in Accounting Research. Journal of Accounting Research 54 (2), 477–523. 10.1111/1475-679X.12116.

Gow, I.D., Ormazabal, G., Taylor, D.J., 2010. Correcting for Cross-Sectional and Time-Series Dependence in Accounting Research. The Accounting Review 85 (2), 483–512. 10.2308/accr.2010.85.2.483.

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

Greiner, A., Kohlbeck, M.J., Smith, T.J., 2017. The Relationship between Aggressive Real Earnings Management and Current and Future Audit Fees. AUDITING: A Journal of Practice & Theory 36 (1), 85–107. 10.2308/ajpt-51516.

Griffin, J.J., Mahon, J.F., 1997. The Corporate Social Performance and Corporate Financial Performance Debate. Business & Society 36 (1), 5–31. 10.1177/000765039703600102.

Gunny, K.A., 2010. The Relation Between Earnings Management Using Real Activities Manipulation and Future Performance: Evidence from Meeting Earnings Benchmarks*. Contemporary Accounting Research 27 (3), 855–888. 10.1111/j.1911-3846.2010.01029.x.

Haans, R.F.J., Pieters, C., He, Z.-L., 2016. Thinking about U: Theorizing and testing U- and inverted U-shaped relationships in strategy research. Strat. Mgmt. J. 37 (7), 1177–1195. 10.1002/smj.2399.

Hamadi, M., Heinen, A., Linder, S., Porumb, V.-A., 2016. Does Basel II affect the market valuation of discretionary loan loss provisions? Journal of Banking & Finance 70, 177–192. 10.1016/j.jbankfin.2016.06.002.

Hasan, I., Wall, L.D., 2004. Determinants of the Loan Loss Allowance: Some Cross-Country Comparisons. The Financial Review 39, 129–152.

Hay, D., Knechel, W.R., Willekens, M. (Eds.), 2014. The Routledge Companion to Auditing. Taylor and Francis Group, London, New York.

Healy, P.M., 1985. The effect of bonus schemes on accounting decisions. Journal of Accounting and Economics 7 (1-3), 85–107. 10.1016/0165-4101(85)90029-1.

Hemingway, C.A., Maclagan, P.W., 2004. Managers' Personal Values as Drivers of Corporate Social Responsibility. J Bus Ethics 50 (1), 33–44. 10.1023/B:BUSI.0000020964.80208.c9.

Hense, B., Ulrich, D. (Eds.), 2013. WPO-Kommentar – Kommentar zum Berufsrecht der Wirtschaftsprüfer und vereidigten Buchprüfer - Wirtschaftsprüferordnung (WPO), 2nd ed. Beck Publishing, Düsseldorf.

Hoechle, D., 2007. Robust standard errors for panel regressions with cross-sectional dependence. The Stata Journal 7 (3), 281–312.

Holthausen, R., Watts, R., 2001. The relevance of the value relevance literature for financial accounting standard setting. Journal of Accounting & Economics 31, 3–75. 10.1016/S0165-4101(01)00029-5.

Hong, Y., Andersen, M.L., 2011. The Relationship Between Corporate Social Responsibility and Earnings Management: An Exploratory Study. J Bus Ethics 104 (4), 461–471. 10.1007/s10551-011-0921-y.

Hoogervorst, H. (2017). How to improve communication in financial reporting, implementation support for IFRS Standards, and the role of the Board in wider corporate reporting. November 09, Speech Delivered at the 14th international CPC accounting seminar in Brazil at Sao Paulo Brazil.

Hope, O.-K., Langli, J.C., 2010. Auditor Independence in a Private Firm and Low Litigation Risk Setting. The Accounting Review 85 (2), 573–605. 10.2308/accr.2010.85.2.573.

Huang, X."B.", Watson, L., 2015. Corporate social responsibility research in accounting. Journal of Accounting Literature 34, 1–16. 10.1016/j.acclit.2015.03.001.

IDW (Ed.), 2005. IDW Rechnungslegungshinweis: Anhangangaben nach § 285 Satz 1 Nr. 17 HGB bzw. § 314 Abs. 1 Nr. 9 HGB über das Abschlussprüferhonorar: IDW RH HFA 1.006.

IDW (Ed.), 2010. IDW Stellungnahme zur Rechnungslegung: Anhangangaben nach §§ 285 Nr. 17, 314 Abs. 1 Nr. 9 HGB über das Abschlussprüferhonorar: IDW RS HFA 36.

IESBA, 2012. Handbook of the Code of Ethics for Professional Accountants. IESBA, New York.

IESBA, 2016. Handbook of the Code of Ethics for Professional Accountants. IESBA, New York.

Jansen, I.P.V.O., Ramnath, S., Yohn, T.L., 2012. A Diagnostic for Earnings Management Using Changes in Asset Turnover and Profit Margin*. Contemporary Accounting Research 29 (1), 221–251. 10.1111/j.1911-3846.2011.01093.x.

Jasch, C. 2015. CSR und Berichterstattung. in: Schneider, A., Schmidtpeter, R. (Eds.), Corporate Social Responsibility – Verantwortungsvolle Unternehmensführung in Theorie und Praxis. 2nd ed. Springer Gabler Publishing, Berlin Heidelberg.

Jensen, M.C., 1993. The Modern Industrial Revolution, Exit, and the Failure of Internal Control Systems. The Journal of Finance 48 (3), 831. 10.2307/2329018.

Jensen, M.C., Meckling, W.H., 1976. Theory of the firm: Managerial behavior, agency costs and ownership structure. Journal of Financial Economics 3 (4), 305–360. 10.1016/0304-405x(76)90026-x.

Jin, J., Kanagaretnam, K., Lobo, G.J., Smith, T., 2016. Discretion in bank loan loss allowance, risk taking and earnings management. Accounting & Finance 77. 10.1111/acfi.12210.

Johnson, R.A., Greening, D.W., 1999. The Effects Of Corporate Governance And Institutional Ownership Types Of Corporate Social Performance. Academy of Management Journal 42 (5), 564–576. 10.2307/256977.

Jones, J.J., 1991. Earnings Management During Import Relief Investigations. Journal of Accounting Research 29 (2), 193. 10.2307/2491047.

Jones, T.M., 1995. Institutional Stakeholder Theory: A Synthesis Of Ethics And Economics. Academy of Management Review 20 (2), 404–437. 10.5465/amr.1995.9507312924.

Jungmann, C., 2006. The Effectiveness of Corporate Governance in One-Tier and Two-Tier Board Systems – Evidence from the UK and Germany –. European Company and Financial Law Review 3 (4). 10.1515/ecfr.2006.019.

Kanagaretnam, K., Krishnan, G.V., Lobo, G.J., 2009. Is the market valuation of banks' loan loss provision conditional on auditor reputation? Journal of Banking & Finance 33 (6), 1039–1047. 10.1016/j.jbankfin.2008.10.013.

Kanagaretnam, K., Krishnan, G.V., Lobo, G.J., 2010a. An Empirical Analysis of Auditor Independence in the Banking Industry. The Accounting Review 85 (6), 2011–2046. 10.2308/accr.2010.85.6.2011.

Kanagaretnam, K., Lim, C.Y., Lobo, G.J., 2010b. Auditor reputation and earnings management: International evidence from the banking industry. Journal of Banking & Finance 34 (10), 2318–2327. 10.1016/j.jbankfin.2010.02.020.

Kanagaretnam, K., Lim, C.Y., Lobo, G.J., 2014. Effects of international institutional factors on earnings quality of banks. Journal of Banking & Finance 39, 87–106. 10.1016/j.jbankfin.2013.11.005.

Kanagaretnam, K., Lobo, G.J., Mathieu, R., 2003. Managerial Incentives for Income Smoothing Through Bank Loan Loss Provisions. Review of Quantitative Finance and Accounting 20, 63–80.

Kanagaretnam, K., Lobo, G.J., Wang, C., 2015. Religiosity and Earnings Management: International Evidence from the Banking Industry. Journal of Business Ethics 132 (2), 277–296. 10.1007/s10551-014-2310-9.

Kanagaretnam, K., Lobo, G.J., Yang, D.-H., 2004. Joint Tests of Signaling and Income Smoothing through Bank Loan Loss Provisions. Contemporary Accounting Research 21 (4), 843–884. 10.1506/UDWQ-R7B1-A684-9ECR.

Kanagaretnam, K., Lobo, G.J., Yang, D.-H., 2005. Determinants of signaling by banks through loan loss provisions. Journal of Business Research 58 (3), 312–320. 10.1016/j.jbusres.2003.06.002.

Kang, S.-H., Sivaramakrishnan, K., 1995. Issues in Testing Earnings Management and an Instrumental Variable Approach. Journal of Accounting Research 33 (2), 353. 10.2307/2491492.

Kilic, E., Lobo, G.J., Ranasinghe, T., Sivaramakrishnan, K., 2013. The Impact of SFAS 133 on Income Smoothing by Banks through Loan Loss Provisions. The Accounting Review 88 (1), 233–260. 10.2308/accr-50264.

Kim, I., Venkatachalam, M., 2011. Are Sin Stocks Paying the Price for Accounting Sins? Journal of Accounting, Auditing & Finance 26 (2), 415–442. 10.1177/0148558x11401222.

Kim, J.-B., Chung, R., Firth, M., 2003. Auditor Conservatism, Asymmetric Monitoring, and Earnings Management*. Contemporary Accounting Research 20 (2), 323–359. 10.1506/J29K-MRUA-0APP-YJ6V.

Kim, M.-S., Kross, W., 1998. The impact of the 1989 change in bank capital standards on loan loss provisions and loan write-offs. Journal of Accounting and Economics 25 (1), 69–99. 10.1016/S0165-4101(98)00015-9.

Kim, Y., Park, M.S., 2005. Pricing of Seasoned Equity Offers and Earnings Management. The Journal of Financial and Quantitative Analysis 40 (02), 435. 10.1017/s0022109000002374.

Kim, Y., Park, M.S., 2014. Real Activities Manipulation and Auditors' Client-Retention Decisions. The Accounting Review 89 (1), 367–401. 10.2308/accr-50586.

Kim, Y., Park, M.S., Wier, B., 2012. Is Earnings Quality Associated with Corporate Social Responsibility? The Accounting Review 87 (3), 761–796. 10.2308/accr-10209.

Kinney, W.R., Palmrose, Z.-V., Scholz, S., 2004. Auditor Independence, Non-Audit Services, and Restatements: Was the U.S. Government Right?*. Journal of Accounting Research 42 (3), 561–588. 10.1111/j.1475-679X.2004.t01-1-00141.x.

Kirsch, H.-J., Ewelt-Knauer, C., Gallasch, F., 2013. Stärkung der wahrgenommenen Unabhängigkeit des Abschlussprüfers durch Angaben zu den Honoraren im Konzernanhang. Zeitschrift für das gesamte Gesellschaftsrecht 42 (5), 647–667.

KLD Research and Analytics, Inc. (KLD), 2006. Getting started with KLD stats and KLD's rating definitions. KLD Research & Analytics, Boston, MA.

Klein, A., 2002. Audit committee, board of director characteristics, and earnings management. Journal of Accounting and Economics 33 (3), 375–400. 10.1016/s0165-4101(02)00059-9.

Klynveld Peat Marwick Goerdeler (KPMG), 2014. Geschäftsberichte lessen und verstehen, 7th ed. Available at https://publicgovernance.de/media/geschaeftsberichte-lesen-und-verstehen-compressed.pdf. Updated on July 2014. Checked on November 29 2017.

Köhler, A.G., Marten, K.-U., Ratzinger, N.V.S., 2010. Prüfungshonorare in Deutschland - Determinanten und Implikationen. Zeitschrift für Betriebswirtschaft 80 (1), 5–29.

Kothari, S.P., Leone, A.J., Wasley, C.E., 2005. Performance matched discretionary accrual measures. Journal of Accounting and Economics 39 (1), 163–197. 10.1016/j.jacceco.2004.11.002.

Kothari, S.P., Ramanna, K., Skinner, D., 2010. Implications for GAAP from an analysis of positive research in accounting. Working Paper: Harvard Business School, MIT, University of Chicago.

Krauß, P., 2008. Publizität von Abschlussprüferhonoraren bei kapitalmarktorientierten Unternehmen - Zielsetzung und Wirkung der Regelungen im Bilanzrechtsreformgesetz. Sternenfels Publishing, Sternenfels.

Laeven, L., Majnoni, G., 2003. Loan loss provisioning and economic slowdowns: Too much, too late? Journal of Financial Intermediation 12 (2), 178–197. 10.1016/S1042-9573(03)00016-0.

Larcker, D.F., Richardson, S.A., 2004. Fees Paid to Audit Firms, Accrual Choices, and Corporate Governance. Journal of Accounting Research 42 (3), 625–658. 10.1111/j.1475-679X.2004.t01-1-00143.x.

Larcker, D.F., Richardson, S.A., Tuna, I., 2007. Corporate Governance, Accounting Outcomes, and Organizational Performance. The Accounting Review 82 (4), 963–1008. 10.2308/accr.2007.82.4.963.

Leffson, U., 1995. Wirtschaftsprüfung, 4th ed. Springer Publishing, Wiesbaden.

Lennox, C., Pittman, J.A., 2010. Big Five Audits and Accounting Fraud*. Contemporary Accounting Research 27 (1), 209–247. 10.1111/j.1911-3846.2010.01007.x.

Lenz, H., 2010. § 319b Netzwerk, in: Petersen, J., Zwirner, C., Brösel, G. (Eds.), Handbuch Bilanzrecht. Beck Publishing, Köln, pp. 117–131.

Lenz, H., Bauer, M., Auerbacher, C., 2006a. Die Unabhängigkeit des Abschlussprüfers aus der Perspektive der ökonomischen und verhaltenswissenschaftlichen Prüfungstheorie am Beispiel gemeinsamer Prüfung und Beratung, in: Meeh, G. (Ed.), Unternehmensbewertung, Rechnungslegung und Prüfung. Schaeffer-Poeschel Publishing, Hamburg, pp. 175–229.

Lenz, H., Möller, M., Höhn, B., 2006b. Offenlegung der Honorare für Abschlussprüferleistungen im Geschäftsjahr 2005 bei DAX-Unternehmen. Betriebs-Berater (33), 1787–1793.

Lepetit, L., Nys, E., Rous, P., Tarazi, A., 2008. Bank income structure and risk: An empirical analysis of European banks. Journal of Banking & Finance 32 (8), 1452–1467. 10.1016/j.jbankfin.2007.12.002.

Leuz, C., Pfaff, D., Hopwood, A. (Eds.), 2004. The Economics and Politics of Accounting. Oxford University Press, Oxford.

Leuz, C., Nanda, D., Wysocki, P.D., 2003. Earnings management and investor protection: An international comparison. Journal of Financial Economics 69 (3), 505–527. 10.1016/s0304-405x(03)00121-1.

Lev, B., 1992. Information Disclosure Strategy. California Management Review 34 (4), 9–32. 10.2307/41166701.

Leventis, S., Dimitropoulos, P.E., Anandarajan, A., 2011. Loan Loss Provisions, Earnings Management and Capital Management under IFRS: The Case of EU Commercial Banks. Journal of Financial Services Research 40 (1-2), 103–122. 10.1007/s10693-010-0096-1.

Levich, R.M., Majnoni, G., Reinhart, C.M. (Eds.), 2002. Ratings, Rating Agencies and the Global Financial System. Springer US, Boston, MA.

Levitt, A. (1998). Numbers Game. September 28, Speech Delivered at New York University, Center For Law And Business.

Li, C., 2009. Does Client Importance Affect Auditor Independence at the Office Level?: Empirical Evidence from Going-Concern Opinions. Contemporary Accounting Research 26 (1), 201–230. 10.1506/car.26.1.7.

Lin, J.W., Hwang, M.I., 2010. Audit Quality, Corporate Governance, and Earnings Management: A Meta-Analysis. International Journal of Auditing 14 (1), 57–77. 10.1111/j.1099-1123.2009.00403.x.

Linthicum, C., Reitenga, A.L., Sanchez, J.M., 2010. Social responsibility and corporate reputation: The case of the Arthur Andersen Enron audit failure. Journal of Accounting and Public Policy 29 (2), 160–176. 10.1016/j.jaccpubpol.2009.10.007.

Liu, C.-C., Ryan, S.G., 1995. The Effect of Bank Loan Portfolio Composition on the Market Reaction to and Anticipation of Loan Loss Provisions. Journal of Accounting Research 33 (1), 77. 10.2307/2491293.

Liu, C.-C., Ryan, S.G., 2006. Income Smoothing over the Business Cycle: Changes in Banks' Coordinated Management of Provisions for Loan Losses and Loan Charge-Offs from the Pre-1990 Bust to the 1990s Boom. The Accounting Review 81 (2), 421–441. 10.2308/accr.2006.81.2.421.

Lobo, G.J., 2017. Accounting research in banking – A review. China Journal of Accounting Research 10 (1), 1–7. 10.1016/j.cjar.2016.09.003.

Lobo, G.J., Yang, D.-H., 2001. Bank Managers' Heterogeneous Decisions on Discretionary Loan Loss Provisions. Review of Quantitative Finance and Accounting 16, 223–250.

Lopatta, K., Kaspereit, T., Canitz, F., Maas, T., 2015. Abschlussprüfung und Honorarabhängigkeit – Eine empirische Untersuchung der Auswirkungen auf die Unabhängigkeit des Abschlussprüfers. Betriebswirtschaftliche Forschung und Praxis (67), 561–591.

Lozano, R., Nummert, B., Ceulemans, K., 2016. Elucidating the relationship between Sustainability Reporting and Organisational Change Management for Sustainability. Journal of Cleaner Production 125, 168–188. 10.1016/j.jclepro.2016.03.021.

Ludewig, R., 2002. Ein Appell - nicht nur an die Wirtschaftsprüfer. Die Wirtschaftsprüfung (55), 613–615.

Ludwig, E. (2009). SEC's Experiment with Reserves Failed. American Banker. 174(54): 9.

Lundholm, R.J., Sloan, R.G., 2004. Equity valuation and analysis. McGraw-Hill / Irwin Publishing Comp., New York.

Mackey, A., Mackey, T.B., Barney, J.B., 2007. Corporate Social Responsibility And Firm Performance: Investor Preferences And Corporate Strategies. Academy of Management Review 32 (3), 817–835. 10.5465/amr.2007.25275676.

Marra, A., Mazzola, P., Prencipe, A., 2011. Board Monitoring and Earnings Management Pre- and Post-IFRS. The International Journal of Accounting 46 (2), 205–230. 10.1016/j.intacc.2011.04.007.

McNichols, M.F., 2000. Research design issues in earnings management studies. Journal of Accounting and Public Policy 19 (4-5), 313–345. 10.1016/S0278-4254(00)00018-1.

McNichols, M.F., 2002. Discussion of the Quality of Accruals and Earnings: The Role of Accrual Estimation Errors. The Accounting Review 77 (s-1), 61–69. 10.2308/accr.2002.77.s-1.61.

McNichols, M.F., Wilson, G.P., 1988. Evidence of Earnings Management from the Provision for Bad Debts. Journal of Accounting Research 26, 1. 10.2307/2491176.

McWilliams, A., Siegel, D., 2001. Corporate Social Responsibility: A Theory Of The Firm Perspective. Academy of Management Review 26 (1), 117–127. 10.5465/amr.2001.4011987.

McWilliams, A., Siegel, D., 2000. Corporate social responsibility and financial performance: Correlation or misspecification? Strat. Mgmt. J. 21 (5), 603–609. 10.1002/(sici)1097-0266(200005)21:5<603:aid-smj101>3.3.co;2-v.

McWilliams, A., Siegel, D.S., Wright, P.M., 2006. Corporate Social Responsibility: Strategic Implications*. J Management Studies 43 (1), 1–18. 10.1111/j.1467-6486.2006.00580.x.

Meeh, G. (Ed.), 2006. Unternehmensbewertung, Rechnungslegung und Prüfung. Schaeffer-Poeschel Publishing, Hamburg.

Menichini, T., Rosati, F., 2014. A Fuzzy Approach to Improve CSR Reporting: An Application to the Global Reporting Initiative Indicators. Procedia - Social and Behavioral Sciences 109, 355–359. 10.1016/j.sbspro.2013.12.471.

Mitra, S., 2016. Nonaudit Service Fees and Auditor Independence: Empirical Evidence from the Oil and Gas Industry. Journal of Accounting, Auditing & Finance 22 (1), 85–107. 10.1177/0148558X0702200106.

Müller, M.A., Riedl, E.J., Sellhorn, T., 2015. Recognition versus disclosure of fair values. The Accounting Review 90 (6), 2411–2447. 10.2308/accr-51044.

Naumann, K.-P., Hamannt, M., 2007. Reform des Berufsrechts der Wirtschaftsprüfer durch das BARefG. Die Wirtschaftsprüfung 60, 901–911.

Naumann, K.-P., 2008. Abschlussprüfung in einem geänderten regulatorischen Umfeld, in: Ballwieser, W., Grewe, W. (Eds.), Wirtschaftsprüfung im Wandel - Herausforderungen an Wirtschaftsprüfung, Steuerberatung, Consulting und Corporate Finance, 1st ed. Beck Publishing, München, pp. 97–120.

O'Brien, R.M., 2007. A Caution Regarding Rules of Thumb for Variance Inflation Factors. Qual Quant 41 (5), 673–690. 10.1007/s11135-006-9018-6.

Ohlson, J., 2000. Residual income valuation: the problems. Working Paper: New York University.

Palepu, K.G., Healy, P.M., Bernard, V.L., 2004. Business analysis & valuation: Using financial statements. South-Western, Cincinnatis, OH.

PCAOB, 2013. Audit Quality Indicators: Standing Advisory Group Meeting Discussion. PCAOB, Washington, DC.

Peasnell, K.V., Pope, P.F., Young, S., 2000. Detecting earnings management using cross-sectional abnormal accruals models. Accounting and Business Research 30 (4), 313–326. 10.1080/00014788.2000.9728949.

Peasnell, K.V., Pope, P.F., Young, S., 2005. Board Monitoring and Earnings Management: Do Outside Directors Influence Abnormal Accruals? J Bus Fin & Acc 32 (7-8), 1311–1346. 10.1111/j.0306-686X.2005.00630.x.

Penman, S., 2009. Financial statement analysis and security valuation, 4th ed. McGraw-Hill Publishing, New York.

Penman, S., Sougiannis, T., 1998. A comparison of dividend, cash flow, and earnings approaches to equity valuation. Contemporary Accounting Research 15, 343–383.

Pérez, D., Salas-Fumás, V., Saurina, J., 2008. Earnings and Capital Management in Alternative Loan Loss Provision Regulatory Regimes. European Accounting Review 17 (3), 423–445. 10.1080/09638180802016742.

Pesaran, M.H., 2004. General diagnostic tests for cross section dependence in panels. Working Paper in Economics 0435, Cambridge, UK: University of Cambridge, Faculty of Economics.

Petersen, J., Zwirner, C., Brösel, G. (Eds.), 2010. Handbuch Bilanzrecht. Beck Publishing, Köln.

Petersen, K., Zwirner, C., 2008. Angabepflicht der Honoraraufwendungen für den Abschlussprüfer - Theoretische und empirische Betrachtung der Offenlegungserfordernisse zur Stärkung der Prüferunabhängigkeit. Die Wirtschaftsprüfung (59), 279–290.

Petersen, M.A., 2008. Estimating Standard Errors in Finance Panel Data Sets: Comparing Approaches. Rev. Financ. Stud. 22 (1), 435–480. 10.1093/rfs/hhn053.

Petrovits, C.M., 2006. Corporate-sponsored foundations and earnings management. Journal of Accounting and Economics 41 (3), 335–362. 10.1016/j.jacceco.2005.12.001.

Pfitzer, N., 2006. Aktuelles zur Qualitätssicherung und Qualitätskontrolle. Die Wirtschaftsprüfung (59), 186–197.

Pfitzer, N., Oser, P., Wader, D., 2007. Der Transparenzbericht gemäß § 55c WPO. WPK-Magazin (4), 54–57.

Porter, M.E., Kramer, M.R., 2006. Strategy and society: The link between competitive advantage and corporate social responsibility. Harvard Business Review 84, 78–92.

Prawitt, D.F., Smith, J.L., Wood, D.A., 2009. Internal Audit Quality and Earnings Management. The Accounting Review 84 (4), 1255–1280. 10.2308/accr.2009.84.4.1255.

Prior, D., Surroca, J., Tribó, J.A., 2008. Are Socially Responsible Managers Really Ethical?: Exploring the Relationship Between Earnings Management and Corporate Social Responsibility. Corporate Governance: An International Review 16 (3), 160–177. 10.1111/j.1467-8683.2008.00678.x.

Pricewaterhouse Coopers (PwC). In depth US2017-24: Contrasting the new US GAAP and IFRS credit impairment models. Accessed 30 November 2017.

Quick, R., Sattler, M., 2011a. Beeinträchtigen Beratungsleistungen die Urteilsfreiheit des Abschlussprüfers? Zum Einfluss von Beratungsleistungen auf die diskretionären Periodenabgrenzungen. Zeitschrift für betriebswirtschaftliche Forschung (63), 310–343.

Quick, R., Sattler, M., 2011b. Das Erfordernis der Umsatzabhängigkeit und die Konzentration auf dem deutschen Markt für Abschlussprüferleistungen. Zeitschrift für Betriebswirtschaft (81), 310–343.

Quick, R., Turley, S., Willekens, M. (Eds.), 2008. Auditing, Trust and Governance. Routledge Companion Hall Publishing, London, New York.

Raftery, A.E., 1995. Bayesian Model Selection in Social Research. Sociological Methodology 25, 111. 10.2307/271063.

Raghunandan, K., Read, W.J., Whisenant, J.S., 2003. Initial Evidence on the Association between Nonaudit Fees and Restated Financial Statements. Accounting Horizons 17 (3), 223–234. 10.2308/acch.2003.17.3.223.

Reynolds, J.K., Francis, J.R., 2001. Does size matter?: The influence of large clients on office-level auditor reporting decisions. Journal of Accounting and Economics 30 (3), 375–400. 10.1016/S0165-4101(01)00010-6.

Reynolds, M., Yuthas, K., 2008. Moral Discourse and Corporate Social Responsibility Reporting. J Bus Ethics 78 (1-2), 47–64. 10.1007/s10551-006-9316-x.

Rogers, W.H., 1993. Regression standard errors in clustered samples. Stata Technical Bulletin 13, 19–23.

Roychowdhury, S., 2006. Earnings management through real activities manipulation. Journal of Accounting and Economics 42 (3), 335–370. 10.1016/j.jacceco.2006.01.002.

Salewski, M., Zülch, H., 2014. The association between corporate social responsibility (CSR) and earnings quality - Evidence from European blue chips. HHL Working Paper 131, Leipzig, Germany: Leipzig Graduate School of Management.

Sattler, M., 2011. Vereinbarkeit von Abschlussprüfung und Beratung. Springer Publishing, Wiesbaden.

Scherer, A.G., Palazzo, G., 2011. The New Political Role of Business in a Globalized World: A Review of a New Perspective on CSR and its Implications for the Firm, Governance, and Democracy. J Management Studies 48 (4), 899–931. 10.1111/j.1467-6486.2010.00950.x.

Schmidt, K., Ebke, W.F. (Eds.), 2008. Münchener Kommentar zum Handelsgesetzbuch, 2nd ed. Beck Publishing, München.

Schneider, A., Schmidtpeter, R. (Eds.), 2015. Corporate Social Responsibility – Verantwortungsvolle Unternehmensführung in Theorie und Praxis. 2nd ed. Springer Gabler Publishing, Berlin Heidelberg.

Schnepel, V., 2013. Kommentierung zu § 55c WPO, in: Hense, B., Ulrich, D. (Eds.), WPO-Kommentar – Kommentar zum Berufsrecht der Wirtschaftsprüfer und vereidigten Buchprüfer - Wirtschaftsprüferordnung (WPO), 2nd ed. Beck Publishing, Düsseldorf.

Schwarz, G., 1978. Estimating the Dimension of a Model. Ann. Statist. 6 (2), 461–464. 10.1214/aos/1176344136.

Securities and Exchange Commission (SEC) (2001). Staff Accounting Bulletin 102: No. 102 - Selected Loan Loss Allowance Methodology and Documentation Issues. July.

Shafer, W.E., 2015. Ethical Climate, Social Responsibility, and Earnings Management. J Bus Ethics 126 (1), 43–60. 10.1007/s10551-013-1989-3.

Sharma, D.S., 2014. Non-audit services and auditor independence, in: Hay, D., Knechel, W.R., Willekens, M. (Eds.), The Routledge Companion to Auditing. Taylor and Francis Group, London, New York, pp. 67–88.

Sharma, V.D., Sharma, D.S., Ananthanarayanan, U., 2011. Client Importance and Earnings Management: The Moderating Role of Audit Committees. AUDITING: A Journal of Practice & Theory 30 (3), 125–156. 10.2308/ajpt-10111.

Shrieves, R.E., Dahl, D., 2003. Discretionary accounting and the behavior of Japanese banks under financial duress. Journal of Banking & Finance 27 (7), 1219–1243. 10.1016/S0378-4266(02)00252-2.

Simunic, D.A., 1984. Auditing, Consulting, and Auditor Independence. Journal of Accounting Research 22 (2), 679. 10.2307/2490671.

Sloan, R.G., 1996. Do Stock Prices Fully Reflect Information in Accruals and Cash Flows about Future Earnings? The Accounting Review 71 (3), 289–315.

Sterling, R., 1970. The Theory of the Measurement of Enterprise Income. Lawrence, Kansas: The University Press of Kansas.

Stickney, C., Brown, P., Wahlen, J., 2004. Financial reporting and statement analysis: A strategic perspective. Thomson South-Western, Cincinnatis, OH.

Stiroh, K.J., 2004. Diversification in Banking: Is Noninterest Income the Answer? Journal of Money, Credit and Banking 36 (5), 853–882.

Stiroh, K.J., Rumble, A., 2006. The dark side of diversification: The case of US financial holding companies. Journal of Banking & Finance 30 (8), 2131–2161. 10.1016/j.jbankfin.2005.04.030.

Subramanyam, K.R., 1996. The pricing of discretionary accruals. Journal of Accounting and Economics 22 (1-3), 249–281. 10.1016/s0165-4101(96)00434-x.

Sultana, A., Willeke, C., 2005. Die neuen Anhangangaben durch das Bilanzrechtsreformgesetz - Darstellung und Erläuterung von § 285 Satz 1 Nr. 17 bis 19 HGB. Steuer und Betrieb (22), 951–957.

Sutton, M. (1997). Current Developments in Financial Reporting: Perspectives From the SEC. Remarks to the AICPA's 1997 National Conference on Banks and Savings Institutions. Washington, DC.

Sutton, R.I., Staw, B.M., 1995. What Theory is Not. Administrative Science Quarterly 40 (3), 371. 10.2307/2393788.

Svanström, T., 2013. Non-audit Services and Audit Quality: Evidence from Private Firms. European Accounting Review 22 (2), 337–366. 10.1080/09638180.2012.706398.

Szwajkowski, E., Figlewicz, R.E., 1999. Evaluating corporate performance: A comparison of the Fortune reputation survey and the Socrates social rating database. Journal of Managerial Issues 11 (2), 137–154.

Tepalagul, N., Lin, L., 2014. Auditor Independence and Audit Quality. Journal of Accounting, Auditing & Finance 30 (1), 101–121. 10.1177/0148558X14544505.

Teoh, S.H., Hwang, C.Y., 1991. Nondisclosure and Adverse Disclosure as Signals of Firm Value. Rev. Financ. Stud. 4 (2), 283–313. 10.1093/rfs/4.2.283.

The Auditing Practices Board, 2010. APB Ethical Standard 4 (Revised) - Fees, Remuneration and Evaluation Policies, Litigation, Gifts and Hospitality: APB ES 4.

Thornton, 2008. Corporate social responsibility: A necessity not a choice. International Business Report, New York.

Trébucq, S., Russ, R., 2005. The case of earnings and stakeholder management: Towards an integrated theory of managerial behaviour. Paper presented at the European Association 28th Annual Congress.

Turban, D.B., Greening, D.W., 1997. Corporate Social Performance And Organizational Attractiveness To Prospective Employees. Academy of Management Journal 40 (3), 658–672. 10.2307/257057.

Turley, S., 2008. Developments in the framework of auditing regulation in the United Kingdom, in: Quick, R., Turley, S., Willekens, M. (Eds.), Auditing, Trust and Governance. Routledge Companion Hall Publishing, London, New York, pp. 205–222.

Van der Meulen, S., Gaeremynck, A., Willekens, M., 2007. Attribute differences between U.S. GAAP and IFRS earnings: An exploratory study. The International Journal of Accounting 42 (2), 123–142. 10.1016/j.intacc.2007.04.001.

Verrecchia, R.E., 1983. Discretionary disclosure. Journal of Accounting and Economics 5, 179–194. 10.1016/0165-4101(83)90011-3.

Verschoor, C.C., 2005. Is there financial value in corporate values? Strategic Finance 87 (1), 17–18.

Victor, B., Cullen, J.B., 1988. The Organizational Bases of Ethical Work Climates. Administrative Science Quarterly 33 (1), 101. 10.2307/2392857.

Vukic, N.M., 2015. Corporate Social Responsibility Reporting: Differences among Selected EU Countries. Business Systems Research Journal 6 (2), 477. 10.1515/bsrj-2015-0012.

Waddock, S.A., Graves, S.B., 1997. The Corporate Social Performance–Financial Performance Link. Strat. Mgmt. J. 18 (4), 303–319. 10.1002/(sici)1097-0266(199704)18:4<303:aid-smj869>3.3.co;2-7.

Wahlen, J.M., 1994. The Nature of Information in Commercial Bank Loan Loss Disclosures. The Accounting Review 69 (3), 455–478.

Warfield, T.D., Wild, J.J., Wild, K.L., 1995. Managerial ownership, accounting choices, and informativeness of earnings. Journal of Accounting and Economics 20 (1), 61–91. 10.1016/0165-4101(94)00393-j.

Watts, R.L., Zimmerman, J.L., 1986. Positive accounting theory. Prentice-Hall, Englewood Cliffs, NJ.

Webb, E., 2004. An Examination of Socially Responsible Firms' Board Structure. Journal of Management & Governance 8 (3), 255–277. 10.1007/s10997-004-1107-0.

Windmeijer, F., 2005. A finite sample correction for the variance of linear efficient two-step GMM estimators. Journal of Econometrics 126 (1), 25–51. 10.1016/j.jeconom.2004.02.005.

Xie, B., Davidson, W.N., DaDalt, P.J., 2003. Earnings management and corporate governance: The role of the board and the audit committee. Journal of Corporate Finance 9 (3), 295–316. 10.1016/s0929-1199(02)00006-8.

Young, S., 1999. Systematic Measurement Error in the Estimation of Discretionary Accruals: An Evaluation of Alternative Modelling Procedures. J Bus Fin & Acc 26 (7&8), 833–862. 10.1111/1468-5957.00277.

Zang, A.Y., 2012. Evidence on the Trade-Off between Real Activities Manipulation and Accrual-Based Earnings Management. The Accounting Review 87 (2), 675–703. 10.2308/accr-10196.

Zhang, J.Q., Zhu, H., Ding, H.-b., 2013. Board Composition and Corporate Social Responsibility: An Empirical Investigation in the Post Sarbanes-Oxley Era. J Bus Ethics 114 (3), 381–392. 10.1007/s10551-012-1352-0.

Zhang, X.F., 2007. Accruals, Investment, and the Accrual Anomaly. The Accounting Review 82 (5), 1333–1363. 10.2308/accr.2007.82.5.1333.

Zimmermann, R.-C., 2006. Gestaltungsspielräume bei Veröffentlichung von Abschlussprüferhonoraren im Rahmen des BilReG. Kapitalmarktorientierte Rechnungslegung (6), 273–276.

Zimmermann, R.-C., 2008. Abschlussprüfer und Bilanzpolitik der Mandanten - eine empirische Analyse des deutschen Prüfungsmarktes, 1st ed. Springer Publishing, Wiesbaden.

Zülch, H., Krauß, P., Pronobis, P., 2010. Die Entwicklung von Abschlussprüferhonoraren in Deutschland zwischen 2004 und 2008: Eine empirische Analyse ausgewählter Börsenindizes (DAX, MDAX, SDAX, TecDAX). Die Wirtschaftsprüfung (63), 397–404.

Zülch, H., Pronobis, P., Krauß, P., 2009. Gefährdung der Unabhängigkeit des Abschlussprüfers durch einen zu hohen Anteil nicht abschlussprüfungsbezogener Leistungen? Zeitschrift für Corporate Governance (4), 226–230.

Appendix

Appendix A – Variable definitions chapter 2

Variable	Definition
	Variables applied in basic models B1-B7
LLP	Total loan loss provisions;
NPA	Non-performing loans/assets;
ΔNPA	Change in non-performing assets;
ALW	Loan loss allowances/reserves;
TL	Total loans;
ΔTL	Change in total loans;
$TL\_CATEGORIES$	Total loan categories, calculated as fraction of loans associated to the respective loan category for each of the categories analysed;
NCO	Net charge-offs, calculated as loan losses minus recoveries;
TA	Total assets.
	Additional variables applied in static and dynamic models S1-S4 and D1-D4
$SIZE$	Bank size, calculated as the natural logarithm of total assets;
$EBTP$	Earnings before tax and provisions, scaled by lagged total assets;
$CAPB$	Capital adequacy ratio, calculated as tier 1 capital, scaled by lagged total assets.
	Control variables
ΔGDP	Change in GDP for country j the respective bank i is located, from year t-1 to year t;

$\Delta UNEMP$	Change in unemployment rate for country j the respective bank i is located, from year t-1 to year t;
$\Delta LandPrice$	Change in land prices, e.g., calculated as change in a land price index;
$CSRET$	Return on the Case-Shiller Real estate Index;
ΔBFI	Change in a business failure index;
ΔSDA	Change in implied standard deviation of bank asset values.
Proxies for earnings management	
$DLLP$ ($DLLP\_abs$)	Discretionary loan loss provisions, comprehended as the signed (absolute) value of the residual from first-stage regressions.
Proxy for extreme performance	
CFO	Cash-flow from operations.
Variables for OMV-test	
$GROWTH$	Change in sales, scaled by lagged total assets;
$LOSS$	Dummy variable, which equals 1 if net income before extraordinary items is negative, 0 otherwise;
$LOANINT$	Loan intensity, calculated as total loans scaled by total assets;
$INCDIV$	Income diversification, calculated as non-interest income scaled by interest income.
Proxy for earnings management detection	
$ARES$	Average residual, calculated as mean value of DLLP for each bank;
$CRKQ$	Comment/restatement of K-10/Q-10 indicator. Dummy variable, which equals 1 if the respective bank received a SEC comment letter on their annual (K-10) or quarterly (Q-10) financial report with respect to the treatment of loan loss provisioning and related

issues; or the respective bank restated its K-10 or Q-10 report because of insufficiencies regarding the treatment of loan provisioning and related issues.

Appendix B – List of specifications chapter 2

No.	*Authors*	*Sample*	*Sample period*	*One vs. two-step models*	*Specification*
Papers used in Beatty and Liao (2014)					
1	**Beatty et al. (1995)** *Journal of Accounting Research*	USA	1985-1989	One	$LLP_{it} = \alpha_0 + \beta_1 NPA_{it} + \beta_2 ALW_{it} + \varepsilon_{it}$
2	**Beaver & Engel (1996)** *Journal of Accounting and Economics*	USA	1977-1991	Two	$LLP_{it} = \alpha_0 + \beta_1 TL_{it} + \beta_2 NPA_{it} + \beta_3 \Delta NPA_{it+1} + \beta_4 NCO_{it} + \varepsilon_{it}$
3	**Beck & Narayanamoorthy (2013)** *Journal of Accounting and Economics*	USA	1992-2008	One	$LLP_{it} = \alpha_0 + \beta_1 \Delta TL_{it} + \beta_{2-3} TL\_CATEGORIES_{it} + \beta_4 NCO_{it} + \beta_5 ALW_{it-1} + \beta_6 SIZE_{it} + \beta_7 CSRET_{it} + \beta_8 \Delta UNEMP_{it} + \varepsilon_{it}$
4	**Bushman & Williams (2012)** *Journal of Accounting and Economics*	International	1995-2006	One	$LLP_{it} = \alpha_0 + \beta_1 NPA_{it-2} + \beta_2 NPA_{it-1} + \beta_3 \Delta NPA_{it} + \beta_4 \Delta NPA_{it+1} + \beta_5 CAPB_{it} + \beta_6 EBTP_{it} + \beta_7 SIZE_{it} + \beta_8 \Delta GDP_{it} + \varepsilon_{it}$
5	**Collins et al. (1995)** *Journal of Accounting Research*	USA	1971-1991	One	$LLP_{it} = \alpha_0 + \beta_1 NPA_{it-1} + \beta_2 \Delta NPA_{it} + \beta_3 ALW_{it-1} + \beta_4 CAPB_{it} + \beta_5 EBTP_{it} + \varepsilon_{it}$
6	**Kanagaretnam et al. (2010a)** *The Accounting Review*	International	2000-2006	Two	$LLP_{it} = \alpha_0 + \beta_1 TL_{it} + \beta_2 \Delta TL_{it} + \beta_{3-4} TL\_CATEGORIES_{it} + \beta_5 NPA_{it-1} + \beta_6 \Delta NPA_{it} + \beta_7 NCO_{it} + \beta_8 ALW_{it-1} + \varepsilon_{it}$
7	**Kim & Kross (1998)** *Journal of Accounting and Economics*	USA	1984-1992	One	$LLP_{it} = \alpha_0 + \beta_1 TL_{it} + \beta_2 \Delta TL_{it} + \beta_3 NPA_{it-1} + \beta_4 \Delta NPA_{it} + \beta_5 NCO_{it} + \beta_6 \Delta ROA_{it} + \beta_7 SIZE_{it} + \varepsilon_{it}$
8	**Liu & Ryan (2006)** *The Accounting Review*	USA	1991-2000	One	$LLP_{it} = \alpha_0 + \beta_1 TL\_CATEGORIES_{it} + \beta_2 \Delta NPA_{it} + \beta_3 CAPB_{it} + \beta_4 EBTP_{it} + \beta_5 Dum(AbMedROA)_{it} + \varepsilon_{it}$

9	**Wahlen (1994)** *The Accounting Review*	USA	1977-1988	Two	$LLP_{it} = \alpha_0 + \beta_1 TL_{it} + \beta_2 NPA_{it-1} + \beta_3 \Delta NPA_{it} + \beta_4 ALW_{it-1} + \varepsilon_{it}$
Other important papers					
10	**Ahmed et al. (1999)** *Journal of Accounting and Economics*	USA	1986-1995	One	$LLP_{it} = \alpha_0 + \beta_1 \Delta NPA_{it} + \beta_2 CAPB_{it} + \beta_3 EBTP_{it} + \beta_4 \Delta BFI_{it} + \beta_5 \Delta SDA_{it} + \varepsilon_{it}$
11	**Agarwal et al. (2007)** *International Review of Economics and Finance*	Japan	1985-1999	One	$LLP_{it} = \alpha_0 + \beta_1 \Delta TL_{it} + \beta_2 ALW_{it-1} + \beta_3 DUM(Bel25QCAP)_{it-1} + \beta_4 DUM(Ab25QBel75QCAP)_{it-1} + \beta_5 DUM(Ab75QCAP)_{it-1} + \beta_6 EBTP_{it} + \beta_7 (EBTP * NEG)_{it} + \beta_8 SIZE_{it-1} + \beta_9 GAINS_{it} + \beta_{10} NETDIV_{it} + \beta_{11} \Delta LandPrice_{it} + \varepsilon_{it}$
12	**Anandarajan et al. (2007)** *Journal of Accounting, Auditing & Finance*	Australia	1991-2001	One	$LLP_{it} = \alpha_0 + \beta_1 \Delta NCO_{it} + \beta_2 CAPB_{it} + \beta_3 EBTP_{it} + \beta_4 SIZE_{it} + \beta_5 NII_{it} + \beta_6 \Delta GDP_{it} + \varepsilon_{it}$
13	**Beatty et al. (2002)** *The Accounting Review*	USA	1988-1998	Two	$LLP_{it} = \alpha_0 + \beta_{1-6} TL\_CATEGOTIES_{it} + \beta_7 \Delta NPA_{it} + \beta_8 ALW_{it} + \beta_9 SIZE_{it} + \varepsilon_{it}$
14	**Beatty & Liao (2011)** *Journal of Accounting and Economics*	USA	1993-2009	Two	$LLP_{it} = \alpha_0 + \beta_1 NPA_{it-2} + \beta_2 NPA_{it-1} + \beta_3 \Delta NPA_{it} + \beta_4 \Delta NPA_{it+1} + \beta_5 CAPB_{it} + \beta_6 EBTP_{it} + \varepsilon_{it}$
15	**Bikker & Metzemakers (2005)** *Journal of International Financial Markets, Institutions and Money*	International	1991-2001	One	$LLP_{it} = \alpha_0 + \beta_1 TL_{it} + \beta_2 \Delta TL_{it} + \beta_3 CAPB_{it} + \beta_4 EBTP_{it} + \beta_5 \Delta GDP_{it} + \beta_6 UNEMP_{it} + \varepsilon_{it}$
16	**Bouvatier et al. (2014)** *Journal of Banking & Finance*	Europe	2004-2009	One	$LLP_{it} = \alpha_0 + \beta_1 LLP_{it-1} + \beta_4 TL_{it} + \beta_5 \Delta TL_{it} + \beta_3 ALW_{it-1} + \beta_4 CAPB_{it-1} + \beta_5 EBTP_{it} + \beta_6 NII_{it} + \beta_7 \Delta GDP_{jt} + \varepsilon_{it}$

17	**Bushman & Williams (2015)** *Journal of Accounting Research*	USA	1996-2009	Two	$LLP_{it} = \alpha_0 + \beta_1 \Delta NPA_{it-2} + \beta_2 \Delta NPA_{it-1} + \beta_3 \Delta NPA_{it} + \beta_4 \Delta NPA_{it+1} + \beta_5 NCO_{it} + \beta_6 CAPB_{it-1} + \beta_7 EBTP_{it} + \beta_8 SIZE_{it} + \varepsilon_{it}$
18	**Cheng et al. (2011)** *Journal of Accounting, Auditing & Finance*	USA	1994-2007	Two	$LLP_{it} = \alpha_0 + \beta_1 \frac{1}{BVE + ALW_{it}} + \beta_2 \Delta TL_{it} + \beta_3 \Delta NPA_{it} + \beta_4 \Delta NPA_{it+1} + \beta_5 NCO_{it} + \varepsilon_{it}$
19	**Cohen et al. (2014)** *Journal of Money, Credit and Banking*	USA	1997-2009	Two	$LLP_{it} = \alpha_0 + \beta_{1-5} TL\_CATEGORIES_{it} + \beta_6 NPA_{it} + \beta_7 ALW_{it} + \beta_8 SIZE_{it} + \varepsilon_{it}$
20	**DeBoskey & Jiang (2012)** *Journal of Banking & Finance*	USA	2002-2006	Two	$LLP_{it} = \alpha_0 + \beta_{1-3} TL\_CATEGORIES_{it} + \beta_4 NPA_{it} + \beta_5 \Delta NPA_{it} + \beta_6 ALW_{it} + \beta_{7-8} CAPB_{it} + \beta_9 EBTP_{it} + \beta_{10} SIZE_{it} + \beta_{11} \Delta Assets_{it} + \varepsilon_{it}$
21	**El Sood (2012)** *International Review of Financial Analysis*	USA	2001-2009	One	$LLP_{it} = \alpha_0 + \beta_1 \Delta TL_{it+1} + \beta_2 \Delta NPA_{it+1} + \beta_3 CAPB_{it} + \beta_4 SIZE_{it} + \beta_5 \Delta STDEQ_{it+1} + \varepsilon_{it}$
22	**Fonseca & Gonzalez (2008)** *Journal of Banking & Finance*	Interna-tional	1995-2002	One	$LLP_{it} = \alpha_0 + \beta_1 \Delta TL_{it} + \beta_2 LLP_{it-2} + \beta_3 LLP_{it-1} + \beta_4 ALW_{it} + \beta_5 CAPB_{it} + \beta_6 EBTP_{it} + \beta_7 \Delta GDP_{it} + \varepsilon_{it}$
23	**Gebhardt & Novotny-Farkas (2011)** *Journal of Business Finance & Accounting*	Interna-tional	2000-2007	One	$LLP_{it} = \alpha_0 + \beta_1 \Delta TL_{it} + \beta_2 NPA_{it-1} + \beta_3 \Delta NPA_{it} + \beta_4 CAPB_{it-1} + \beta_5 EBTP_{it} + \varepsilon_{it}$
24	**Hamadi et al. (2016)** *Journal of Banking & Finance*	Interna-tional	2006-2011	Two	$LLP_{it} = \alpha_0 + \beta_1 TL_{it} + \beta_2 \Delta TL_{it} + \beta_3 NPA_{it} + \beta_4 \Delta NPA_{it} + \beta_5 NCO_{it} + \beta_6 CAPB_{it} + \beta_7 SIZE_{it} + \beta_8 \Delta GDP_{it} + \beta_9 \Delta UNEMP_{it} + \beta_{10} HPI_{it} + \beta_{11} TermSpread_{it} + \varepsilon_{it}$
25	**Hasan & Wall (2004)** *The Financial Review*	USA, Inter-national, Ja-pan, Canada	1993-2000	One	$LLP_{it} = \alpha_0 + \beta_1 TL_{it} + \beta_2 NPA_{it} + \beta_3 NCO_{it} + \beta_4 CAPB_{it-1} + \beta_5 EBTP_{it} + Year + \varepsilon_{it}$

26	**Kanagaretnam et al. (2003)** *Review of Quantitative Finance and Accounting*	USA	1987-2000	Two	$LLP_{it} = \alpha_0 + \beta_1 \Delta TL_{it} + \beta_2 NPA_{it-1} + \beta_3 \Delta NPA_{it} + \varepsilon_{it}$
27	**Kanagaretnam et al. (2004)** *Contemporary Accounting Research*	USA	1980-1997	One	$LLP_{it} = \alpha_0 + \beta_1 \Delta TL_{it} + \beta_2 NPA_{it-1} + \beta_3 \Delta NPA_{it} + \beta_4 NCO_{it} + \beta_5 ALW_{it-1} + \beta_6 EBTP_{it} + \varepsilon_{it}$
28	**Kanagaretnam et al. (2005)** *Journal of Business Research*	USA	1980-1997	Two	$LLP_{it} = \alpha_0 + \beta_1 \Delta NPA_{it} + \beta_2 NCO_{it} + \beta_3 ALW_{it-1} + \beta_4 CAPB_{it} + \beta_5 EBTP_{it} + \beta_6 \Delta EBTP_{it+1} + \varepsilon_{it}$
29	**Kanagaretnam et al. (2009)** *Journal of Banking & Finance*	USA	1993-2004	Two	$LLP_{it} = \alpha_0 + \beta_1 \Delta TL_{it} + \beta_2 NPA_{it-1} + \beta_3 \Delta NPA_{it} + \beta_4 NCO_{it} + \beta_5 ALW_{it-1} + \beta_6 CAPB_{it} + \beta_7 EBTP_{it} + Year + \varepsilon_{it}$
30	**Kanagaretnam et al. (2010b)** *Journal of Banking & Finance*	International	1993-2006	Two	$LLP_{it} = \alpha_0 + \beta_1 TL_{it} + \beta_2 \Delta TL_{it} + \beta_3 TL\_CATEGORIES_{it} + \beta_4 NPA_{it} + \beta_5 \Delta NPA_{it} + \beta_6 NCO_{it} + \beta_7 ALW_{it-1} + \varepsilon_{it}$
31	**Kanagaretnam et al. (2014)** *Journal of Banking & Finance*	International	1993-2006	Two	$LLP_{it} = \alpha_0 + \beta_1 TL_{it} + \beta_2 \Delta TL_{it} + \beta_3 TL\_CATEGORIES_{it} + \beta_4 NPA_{it} + \beta_5 NCO_{it} + \beta_6 ALW_{it-1} + Year + Country + \varepsilon_{it}$
32	**Kanagaretnam et al. (2015)** *Journal of Business Ethics*	International	1995-2006	Two	$LLP_{it} = \alpha_0 + \beta_1 TL_{it} + \beta_2 \Delta TL_{it} + \beta_3 TL\_CATEGORIES_{it} + \beta_4 NPA_{it-1} + \beta_5 \Delta NPA_{it-1} + \beta_6 NCO_{it} + Year + Country + \varepsilon_{it}$
33	**Kilic et al. (2013)** *The Accounting Review*	USA	1998-2003	One	$LLP_{it} = \alpha_0 + \beta_1 TL_{it-1} + \beta_2 \Delta TL_{it} + \beta_3 NPA_{it-1} + \beta_4 \Delta NPA_{it} + \beta_5 NCO_{it} + \beta_6 ALW_{it} + \beta_7 CAPB_{it} + \beta_8 \Delta EBTP_{it} + \beta_9 Dum(Ab75QBel25QEBTP)_{it} + \varepsilon_{it}$
34	**Laeven & Majnoni (2003)** *Journal of Financial Intermediation*	International	1988-1999	One	$LLP_{it} = \alpha_0 + \beta_1 \Delta TL_{it} + \beta_2 EBTP_{it} + \beta_3 \Delta GDP_{it} + \varepsilon_{it}$
35	**Leventis et al. (2011)** *Journal of Financial Services Research*	Europe	1999-2008	One	$LLP_{it} = \alpha_0 + \beta_1 CAPB_{it} + \beta_2 EBTP_{it} + \beta_3 SIZE_{it} + \beta_4 NII_{it} + \beta_5 \Delta GDP_{it} + \varepsilon_{it}$

36	**Liu & Ryan (1995)** *Journal of Accounting Research*	USA	1983-1991	One	$LLP_{it} = \alpha_0 + \beta_1 \Delta NPA_{it-3} + \beta_2 \Delta NPA_{it-2} + \beta_3 \Delta NPA_{it-1} + \beta_4 \Delta NPA_{it} + \varepsilon_{it}$
37	**Lobo & Yang (2001)** *Review of Quantitative Finance and Accounting*	USA	1981-1996	One	$LLP_{it} = \alpha_0 + \beta_1 \Delta TL_{it} + \beta_2 NPA_{it-1} + \beta_3 \Delta NPA_{it} + \beta_4 NCO_{it} + \beta_5 ALW_{it-1} + \beta_6 CAPB_{it} + \beta_7 EBTP_{it} + \varepsilon_{it}$
38	**Pérez et al. (2008)** *European Accounting Review*	Spain	1986-2000	One	$LLP_{it} = \alpha_0 + \beta_1 TL_{it} + \beta_2 NPA_{it} + \beta_3 CAPB_{it-1} + \beta_4 EBTP_{it} + \beta_5 SIZE_{it} + \beta_6 \Delta GDP_t + \beta_7 IBOL_t + \varepsilon_{it}$
39	**Shrieves & Dahl (2003)** *Journal of Banking & Finance*	Japan	1989-1996	One	$LLP_{it} = \alpha_0 + \beta_1 \Delta TL_{it} + \beta_2 ALW_{it-1} + \beta_3 DUM(Bel25QCAP)_{it-1} + \beta_4 DUM(Ab25QBel75QCAP)_{it-1} + \beta_5 DUM(Ab75QCAP)_{it-1} + \beta_6 EBTP_{it} + \beta_7 (EBTP * NEG)_{it} + \beta_8 SIZE_{it-1} + \beta_9 Liabilities_{it} + \beta_{10} GAINS_{it} + \beta_{11} NETDIV_{it} + \beta_{12} \Delta LandPrice_{it} + \varepsilon_{it}$

Appendix C – Variable definitions chapter 3

Variable Definitions

Variable	Definition
$DACC\_abs$	Absolute value of discretionary accruals, calculated as residuals from the cross-sectional McNichols model;
$DACC$	Signed value of discretionary accruals, calculated as residuals from the cross-sectional McNichols model. $PDACC$ contains only positive, $NDACC$ only negative values;
REM	Signed value of the sum of real activities based EM measures, calculated as negatively scaled proxy;
EM_ALL	Total EM proxy, calculated as sum of deciles of both $DACC\_abs$ and REM;
$REMvsAEM$	EM trade-off proxy, calculated as ratio of REM decile and EM_ALL;
$Q1DACC\_abs$	Extreme AEM indicator, calculated as 1 if $DACC\_abs$ is in the top quintile of the distribution, 0 otherwise;
$Q1REM$	Extreme REM indicator, calculated as 1 if inverted REM is in the top quintile of the distribution, 0 otherwise.
CI	Client importance, either proxied by CIS or $CINAS$;
CIS	Client Importance Sales, calculated as share of total fees for audit and non-audit services of the total sales of the audit firm in financial year. Total sales of the audit firm have to disclosed in the transparency reports for each financial year according to Art. 55c para. 1 sent. 3 no. 3 WPO. Total fees for audit and non-audit services consists of four different numbers, fees from audit services (Art. 314 para. 1 no. 9 HGB), fees from other assurance services (Art. 314 para. 1 no. 9 HGB), fees from tax services (Art. 314 para. 1 no. 9 HGB) and fees from other services (Art. 314 para. 1 no. 9 HGB);
$CINAS$	Client Importance Non-Audit Sales, calculated as share of total non-audit fees of the total fees for audit and non-audit services in financial year. Total non-audit fees consist fees from other assurance services (Art. 314 para. 1 no. 9 HGB), fees from tax services (Art. 314 para. 1 no. 9 HGB) and fees from other services (Art. 314 para. 1 no. 9 HGB). Total fees for audit and non-audit services consists of all four different numbers, fees from audit services (Art. 314 para. 1 no. 9 HGB), fees from other assurance services (Art. 314 para. 1 no. 9 HGB), fees from tax services (Art. 314 para. 1 no. 9 HGB) and fees from other services (Art. 314 para. 1 no. 9 HGB);
DEP	Dependent Firm; indicator, which is 1 when firm's respective client importance proxy lies above the discussed thresholds (5%, 10%, 15% for CIS, 40%, 50% for $CINAS$), 0 otherwise. See detailed information on whether the dependence is studied for a one, two or three consecutive years time-period;
$Ratio\ A$	Importance of the audit service line sales for the total sales of the auditor, calculated as the sum of the audit fees of the auditor in the financial year divided by the total sales of the auditor;
$Ratio\ B$	Importance of total fees from all PIEs for the total sales of the auditor, calculated as the sum of all fees from PIE clients of the auditor in the financial year divided by the total sales of the auditor;

$ASIL$	Low audit service line importance indicator, calculated as 1 when $Ratio\ A$ is below median $Ratio\ A$, 0 otherwise;
$TFIPL$	Low total fees from PIEs importance indicator, calculated as 1 when $Ratio\ B$ is below median $Ratio\ B$, 0 otherwise;
Aud_Mid	EU SME definition from May 2003: midget audit firm ≤ 2 m. EUR Sales;
Aud_Small	EU SME definition from May 2003: small audit firm 2 mi. EUR < sales ≤ 10 m. EUR;
Aud_Medium	EU SME definition from May 2003: medium audit firm 10 m. EUR < sales ≤ 50 m. EUR;
Aud_Large	EU SME definition from May 2003: large audit firm sales > 50 m. EUR;
$AudSize$	Auditor Size; categorical variable, 1 = Aud_Mid; 2 = Aud_Small; 3 = Aud_Medium; 4 = Aud_Large;
AC	Best practice audit committee indicator, calculated as 1 if firm coincides with DCGK paragraph 5.3.2, e.g., existence of an audit committee with a chair that meets the following requirements: (1) independent member of the board, (2) accounting expert, (3) no executive member for at least three years, (4) not the chair of the supervisory board.
ACC	Total accruals, calculated as earnings before extraordinary items minus cash flow from operations;
REV	Revenues;
ΔREV	Change in revenues;
ΔAR	Change in accounts receivable;
PPE	Net property, plant and equipment;
ΔGDP	Change in gross domestic product;
ROA	Return on assets, given as net income before extraordinary items scaled by lagged total assets;
ΔINV	Change in inventories;
$COGS$	Costs of goods sold;
$PROD$	Production costs, calculated as sum of $COGS$ sold and ΔINV;
$DISEXP$	Discretionary expenses, calculated as sum of R&D and selling, general and administrative expenses (Advertising expenses also included);
$Tobin'sQ$	Tobin's Q, calculated as sum of MVE, long-term debt and current debt, divided by total assets;
INT	Internal funds, calculated as sum of net income before extraordinary items, research and development expenses and depreciation and amortization.
$Size$	Natural logarithm of the market value of equity (MVE);
MB	Market-to-book ratio, calculated as MVE/BVE, where BVE is the book value of equity;
$LOSS$	Indicator, which is 1 when firm's net income before extraordinary items is negative, and 0 otherwise;

$Litigation$	Indicator for litigation risk of firms in regulated markets, which is 1 when firm is located in a market with SIC Code 2833-2836, 3570-3577, 3600-3674, 5200-5961 or 7370-7474, and 0 otherwise;
CFO	Net cash flow from operations;
$Leverage$	Leverage of the firm, given as total debt scaled by total assets;
ΔTA	Change in total assets, scaled by lagged total assets;
$\Delta EARN$	Change in earnings, scaled by lagged total assets;
$FirmAge$	Firm age, calculated as natural logarithm of (1 + the number of years since the firm appears in Thomson Reuters data base);
$PartChange$	Partner Change; indicator, which is 1 when there is a change in audit mandate partners, meaning either engagement or revenue partner or both, and 0 otherwise;
$AudChange$	Indicator, which is 1 when there is a change in the audit firm, and 0 otherwise;
$AudOpClean$	Clean audit opinion indicator, calculated as 1 if the auditor issues a clean audit opinion, 0 otherwise;
$BoardSize$	Board size, calculated as number of independent and executive members on the board;
$BoardInd$	Board Independence, calculated as percentage of independent, meaning non-executive board members on the board;
$VioDCGK$	Numbers of violations on the German Corporate Governance Codex (Deutscher Corporate Governance Kodex); Art. 161 of the German Stock Corporation Act (Aktiengesetz) obliges every capital market oriented German corporation to declare compliance with codex, while the variable captures statement on probable violations;

Appendix D – Variable definitions chapter 4

Variable Definitions

Variable	Definition
$DACC\_abs$	Absolute value of discretionary accruals, calculated as residuals from the cross-sectional McNichols model;
$DACC$	Signed value of discretionary accruals, calculated as residuals from the cross-sectional McNichols model. $PDACC$ contains only positive, $NDACC$ only negative values;
REM_abs	Absolute value of the sum of real activities based EM measures, calculated as negatively scaled proxy: $CFO\_abn_{it} - PROD\_abn_{it} + DISEXP\_abn_{it}$;
REM	Signed value of the sum of real activities based EM measures, calculated as negatively scaled proxy.
EM_ALL	Total EM proxy, calculated as sum of deciles of both $DACC\_abs$ and REM.
$REMvsAEM$	EM trade-off proxy, calculated as ratio of REM decile and EM_ALL.
CSR_Score	Net performance score for the Corporate Social Responsibility pillar from the ASSET4 ESG data base, calculated as average score of the KPI categories assuming equal weights and divided by 100, see Appendix B;
CSR	Indicator, which is 1 when firm's CSR performance score in the year lies above the median of the CSR performances scores of the full sample, and zero otherwise;
REP	Indicator, which is 1 when a firm provides a CSR report, which has been prepared following G4 GRI guidelines for sustainability reporting, and zero otherwise.
BI	Board Independence, calculated as percentage of independent board members, e.g. members that have not served for more than ten years, do have holdings of less than 5 percent, do not have cross-board membership, do not have family ties to the firm and do not accept any compensation other than the compensation for the board service;
ACC	Total accruals, calculated as earnings before extraordinary items minus cash flow from operations, scaled by lagged total assets;
$SCALE$	Scaled intercept, calculated as one divided by lagged total assets;
CFO	Net cash flow from operations, scaled by lagged total assets;
REV	Revenues, scaled by lagged total assets;
ΔREV	Change in revenues, scaled by lagged total assets;
ΔAR	Change in accounts receivable, scaled by lagged total assets;
PPE	Net property, plant and equipment, scaled by lagged total assets;
ΔGDP	Change in gross domestic product, scaled by lagged gross domestic product;
ROA	Return on assets, given as net income before extraordinary items scaled by lagged total assets;
ΔINV	Change in inventories, scaled by lagged total assets;

$COGS$	Costs of goods sold, scaled by lagged total assets;
$PROD$	Production costs, calculated as sum of costs of goods sold and change in inventories, scaled by lagged total assets;
$DISEXP$	Discretionary expenses, calculated as sum of R&D and selling, general and administrative expenses (Advertising expenses also included), scaled by lagged total assets;
$Size$	Natural logarithm of the market value of equity (MVE);
MB	Market-to-book ratio, calculated as MVE/BVE, where BVE is the book value of equity;
$Loss$	Indicator, which is 1 when firm's net income before extraordinary items is negative, and 0 otherwise;
$Leverage$	Leverage of the firm, given as total debt scaled by total assets;
$\Delta EMPLOY$	Change in number of employees, scaled by lagged total assets;
ΔTA	Change in total assets, scaled by lagged total assets;
GOV_Score	Net performance score for the Corporate Governance pillar from the ASSET4 ESG data base, calculated as average score of the KPI categories assuming equal weights and divided by 100. When studying ACI, we leave out the first KPI category (Board Functions) when calculating the net performance score. When studying BI, we leave out the second KPI category (Board Structure) when calculating the net performance score. See Appendix B;
ATO	Asset turnover ratio, calculated as level of revenues divided by net operating assets;
$Litigation$	Indicator for litigation risk of firms in regulated markets, which is 1 when firm is located in a market with SIC Code 2833-2836, 3570-3577, 3600-3674, 5200-5961 or 7370-7474, and 0 otherwise;
$FirmAge$	Firm age, calculated as natural logarithm of (1 + the number of years since the firm appears in Thomson Reuters data base); and
$Risk$	Market risk, given as beta factor of the firm, calculated as relationship of the volatility of the stock of the firm and the volatility of the market, based on between 23 and 35 consecutive month end price changes in percent relative to a local market index.
$SUSPECT$	Indicator for Suspect firms. For the zero earnings threshold, suspect firms are firms within the earnings interval [0;0.005), where intervals are drawn upon net income before extraordinary items scaled by lagged total assets. For the forecasted earnings threshold, suspect firms are firms with an analyst forecast error in the error interval [0;0.005), where intervals are drawn upon error in reported earnings per share compared to mean forecasted earnings per share scaled by mean forecasted earnings per share.

Appendix E – CSR & CG performance score calculation chapter 4

Within the Thomson Reuters ASSET4 ESG data, the CSR_Score_{it} uses the category performance scores of the two pillars Environmental and Social performance:

CSR performance Thomson Reuters ASSET4 ESG		
Pillar	**Environmental**	**Social**
KPI Categories	Emission Reduction Product Innovation Resource Reduction	Customer / Product Responsibility Society / Community Society / Human Rights Workforce / Diversity and Opportunity Workforce / Employment Quality Workforce / Health & Safety Workforce / Training and Development

The CG score GOV_Score_{it} is calculated using the following KPI categories:

CG Performance Thomson Reuters ASSET4 ESG	
Pillar	**Corporate Governance**
KPI Categories	Board of Directors/ Board Functions Board of Directors/ Board Structure Board of Directors/ Compensation Policy Integration/ Vision and Strategy Shareholders / Shareholder Rights

Curriculum Vitae

Stand März 2018

Daniel Schaupp (M.Sc.)

Wissenschaftlicher Mitarbeiter *Lehrstuhl für BWL, insbesondere Controlling und Interne Unternehmensrechnung*
Julius-Maximilians-Universität (JMU) Würzburg
daniel.schaupp@uni-wuerzburg.de

AKADEMISCHE UND SCHULISCHE AUSBILDUNG

Akademische Positionen

Okt 2014 – Mär 2018	***Wissenschaftlicher Mitarbeiter*** *Lehrstuhl für BWL, insbesondere Controlling und Interne Unternehmensrechnung (Prof. Dr. Andrea Szczesny)* *Julius-Maximilians-Universität (JMU) Würzburg, Deutschland* ***Dissertation****: Earnings Management in the Context of Earnings Quality, Corporate Governance and Corporate Social Responsibility*
Aug 2017 – Nov 2017	***Visiting Ph.D.*** *Department of Accounting (Prof. Margaret Abernethy)* *The University of Melbourne, Australien*

Akademische Ausbildung

Apr 2012 – Aug 2014	***Master of Science*** *Business Management JMU Würzburg, Deutschland* *Schwerpunkte: Finance, Accounting, Controlling, Taxation, Econometrics* *Thesis: Earnings management in the context of executive turnovers – empirical analysis of the German HDAX*
Sep 2010 – Jan 2011	***Exchange Term*** *Department of Management & Business* *Prifysgol Aberystwyth University, Wales (UK)* *Schwerpunkte: Corporate Finance, Financial & Management Accounting*
Okt 2008 – Feb 2012	***Bachelor of Science*** *Wirtschaftswissenschaften JMU Würzburg* *Schwerpunkte: Finance, Accounting, Auditing, Taxation* *Thesis: Accounting and configuration possibilities of stock option plans – empirical analysis of the German DAX30*
Jun 2008	***Allgemeine Hochschulreife*** *Frobenius Gymnasium Hammelburg, Deutschland*

WISSENSCHAFTLICHE ARBEITEN

Working Papers

Schaupp (2016). Corporate Social Responsibility and Earnings Management – A Differentiated View on CSR Incentives and the Role of Corporate Governance.

Schaupp, Lenz (2017). Does Client Importance Influence Accrual and Real Activities Earnings Management in the Small and Midsized Audit Market for Listed Clients?

Schaupp, Stralla (2017). Earnings Management Modelling in the Banking Industry – Evaluating Valuable Approaches.

KONFERENZPRÄSENTATIONEN & WORKSHOPS

European Accounting Association Annual Congress

Maastricht, Niederlande 11-13 Mai 2016

Schaupp (2016). The Influence of Corporate Social Responsibility and Board Characteristics on Earnings Management – European Evidence.

Valencia, Spanien 10-12 Mai 2017

Schaupp, Lenz (2017). Does Client Importance Influence Accrual and Real Activities Earnings Management in the Small and Midsized Audit Market for Listed Clients?

Schaupp, Stralla (2017). Earnings Management Modelling in the Banking Industry – Evaluating Valuable Approaches.

Chair in Session FRPS12 (Financial Reporting of Goodwills)

German Academic Association for Business Research (VHB) Conference

München, Deutschland 18-20 Mai 2016

Schaupp (2016). The Influence of Corporate Social Responsibility and Board Characteristics on Earnings Management.

Workshop on Corporate Governance

Mailand, Italien 27-28 Oktober 2016

Schaupp (2016). Corporate Social Responsibility and Earnings Management – A Differentiated View on CSR Incentives and the Role of Corporate Governance.

International Conference on Accounting and Finance

Tokyo, Japan 28-29 Mai 2017

Schaupp, Stralla (2017). Earnings Management Modelling in the Banking Industry – Evaluation Valuable Approaches.

British Accounting and Finance Association Annual Conference

Edinburgh, UK 10-12 April 2017

Schaupp, Lenz (2017). Does Client Importance Influence Accrual and Real Activities Earnings Management in the Small and Midsized Audit Market for Listed Clients?