IFRS, enforcement, and their role for accounting quality and comparability

– An empirical analysis –

Leonard La Roche* and Jan Schotte†

Master Thesis in Accounting and Financial Management at the Stockholm School of Economics

Abstract

Whether IFRS have been associated with an increase in accounting quality and comparability is a debated topic. Prior research mostly shows positive capital market effects (Daske *et al.*, 2011; Li, 2010). However, whether the sources of these effects were increases in accounting quality and comparability has not been tested directly but rather been assumed. Moreover, these effects are often limited to countries with certain institutional characteristics such as strong enforcement regulation. Less evidence exists on the first-order effects of IFRS adoption and enforcement regulation on accounting quality and comparability, i.e. the asserted channels. For this reason, our thesis examines the effects of IFRS adoption and enforcement on accounting quality and comparability. In this context, we analyze a twelve-year time period using a dataset consisting of 24 countries. While accounting quality is measured in terms of earnings management and value relevance, for accounting comparability we use a measure recently developed by De Franco *et al.* (2011). Contrary to IFRS' objectives, we find that both accounting quality and comparability decrease following IFRS adoption. Furthermore, our findings indicate that the importance of concurrent enforcement changes seems to be limited to the second-order capital market effects documented by prior research. Thus, our thesis contributes to existing literature by extending the knowledge of the influence of IFRS and enforcement on reporting practices. Further, we identify potential shortcomings of the accounting comparability measure.

Keywords: accounting harmonization, accounting comparability, accounting quality, earnings management, IFRS, enforcement, reporting incentives

Tutor: Katerina Hellström
Date: 15-05-2013

*40326@student.hhs.se
†40325@student.hhs.se

The authors would like to thank Katerina Hellström for her valuable guidance throughout the whole thesis project, Per-Olov Edlund for his support in statistical matters, and Matthias Breuer for the provision of computer resources as well as help in STATA questions.

Table of Contents

1 Introduction 1

2 Background 3

2.1 Motivation of IFRS adoption 3

2.2 Differentiation of first- and second-order effects (of IFRS adoption) 4

2.3 Empirical evidence on second-order effects 6

2.3.1 Empirical evidence on voluntary IFRS adoption 6

2.3.2 Empirical evidence on mandatory IFRS adoption 7

2.4 Institutional setting and the empirical evidence on first-order effects 9

2.4.1 Definition of the institutional setting 9

2.4.2 Influence of the institutional setting on accounting quality 10

2.4.3 Influence of the institutional setting on accounting comparability 15

2.4.4 Institutional complementarities 17

2.5 Conclusion and implications for our thesis 18

3 Hypotheses development 19

3.1 Hypotheses concerning accounting quality 19

3.2 Hypotheses concerning accounting comparability 20

4 Research design 21

4.1 Sample Selection 21

4.2 Clustering of countries 21

4.3 Time periods 24

4.4 Accounting quality 25

4.4.1 Earnings management 26

4.4.2 Value relevance 30

4.5 Accounting comparability 31

4.5.1 The comparability measure of De Franco *et al.* (2011) 31

4.5.2 Test design 33

5 Data 34

6 Empirical Results 37

6.1 Accounting Quality 37

6.1.1 Description of results 38

6.1.2 Evaluation of hypotheses 47
6.2 Accounting comparability 47
6.2.1 Description of results 47
6.2.2 Evaluation of hypotheses 52
7 Discussion 53
7.1 Accounting Quality 53
7.1.1 Comparison with prior research 53
7.1.2 The role of enforcement for accounting quality 54
7.1.3 Further implications 55
7.2 Accounting Comparability 58
7.2.1 Discrepancy between the two metrics 58
7.2.2 Further implications 59
7.3 Robustness tests 61
7.3.1 The case of the UK 61
7.3.2 Subsequent enforcement reforms 64
7.3.3 Different IFRS adoption dates 65
7.4 Limitations and further areas of future research 67
8 Conclusion 68
References 71
Appendix 77

List of Tables

Table 1: Results of k-means cluster analysis 23
Table 2: Country clusters 24
Table 3: Number of firm-year observations by period and cluster for each panel 35
Table 4: Number of firm-year observations by country for each panel 36
Table 5: Metric 1.1 Variance of *ΔNI** (Panel 1) 39
Table 6: Metric 1.2 Variance of *ΔNI*/ΔCF** (Panel 2) 39
Table 7: Metric 1.3 Correlation between accruals and cash flows (Panel 3) 40
Table 8: Metric 1.4 Small positive earnings (Panel 4) 40
Table 9: Metric 1.5 Large negative net income (Panel 5) 41
Table 10: Metric 2.1 Regression of P* - *adjusted R²* (Panel 6) 45
Table 11: Metric 2.2 Regression of (NI/P)* - *adjusted R²* good news (Panel 6) 46
Table 12: Metric 2.3 Regression of (NI/P)* - *adjusted R²* bad news (Panel 6) 46
Table 13: Metric 3.1 *AGG.ACOMP* (Panel 7) 49
Table 14: Metric 3.2 *AGG.CFCOMP* (Panel 7) 51

List of Figures

Figure 1: Categorization of the effects of IFRS adoption 5
Figure 2: Determinants of accounting quality (Soderstrom and Sun, 2007) 10
Figure 3: Accounting quality in terms of earnings management 44
Figure 4: Return-based accounting comparability metric 49
Figure 5: Cash flow-based accounting comparability metric 51
Figure 6: Accounting quality in terms of earnings management excluding the UK 62
Figure 7: Return-based accounting comparability metric excluding the UK 63
Figure 8: Cash flow-based accounting comparability metric excluding the UK 64
Figure 9: Accounting quality excluding HK, SE, TR 65
Figure 10: Accounting quality excluding IL, NZ, PK, SG, TR 66
Figure 11: Return-based acc. comparability metric excluding IL, NZ, PK, SG, TR 67
Figure 12: Cash flow-based acc. comparability metric excluding IL, NZ, PK, SG, TR 67

List of Appendices

Appendix 1: Countries which adopted IFRS before 2009 77
Appendix 2: Sample countries 78
Appendix 3: Description of variables of institutional characteristics 79
Appendix 4: Institutional characteristics (Djankov *et al.*, 2008; La Porta *et al.*, 2006) 80
Appendix 5: Metrics of accounting quality 81
Appendix 6: Accounting comparability metric based on De Franco *et al.* (2011) 83
Appendix 7: Descriptive statistics for variables used for the accounting quality metrics 85
Appendix 8: Robustness test excluding the UK 87
Appendix 9: Robustness test excluding HK, SE, TR 91
Appendix 10: Robustness test excluding IL, NZ, PK, SG, TR 94

List of Abbreviations

ERC – Enforcement reform countries

EU – European Union

GAAP – Generally Accepted Accounting Principles

HK – Hong Kong

IAS – International Accounting Standards

IASB – International Accounting Standards Board

IASC – International Accounting Standards Committee

IFRS – International Financial Reporting Standards

IL – Israel

NZ – New Zealand

PK – Pakistan

SE – Sweden

SEC – Strong enforcement countries

SG – Singapore

TR – Turkey

UK – United Kingdom

WEC – Weak enforcement countries

1 Introduction

In times of increasing globalization and highly international capital markets, the need for a single set of consistent high quality financial reporting standards gained widespread acceptance among policy makers, standard setters and preparers (e.g. Ball, 1995; Godfrey *et al.*, 2010; Whittington, 2005). In this context, the introduction of International Financial Reporting Standards (IFRS) by a large number of countries presents a historic step and has been motivated by the aim of establishing one set of high quality financial reporting standards fostering the comparability of financial statements across countries. Theoretically, the effects of IFRS adoption can be differentiated into two kinds of effects: changing the accounting standards might lead to direct *first-order effects* on the quality and comparability of financial statements which in turn can lead to *second-order effects* on capital markets. Conceptually, proponents of IFRS argue that the capital market-oriented character of the standards increases accounting quality and comparability, leading to positive second-order effects such as higher market efficiency, lower cost of capital for firms, and a lower investors' risk. (e.g. Ball, 2006; Hail *et al.*, 2010). Other authors, however, acknowledge that due to the principle-based nature of IFRS considerable discretion exists and will be used differently across countries which might speak against an expected increase in terms of accounting quality and comparability (Barth *et al.*, 2008; Ormrod and Taylor, 2004).

Prior research mostly finds that positive second-order effects following IFRS adoption are limited to firms from countries with particular institutional characteristics such as strong enforcement regimes or high governance quality. Concerning first-order effects, empirical studies show inconsistent and sometimes conflicting results. Generally though, prior studies underline the importance of institutional factors for the occurrence of both first-order (Ball *et al.*, 2003; Samarasekera *et al.*, 2012; Soderstrom and Sun, 2007) and second-order effects (Daske *et al.*, 2008; Li, 2010), implying that harmonizing accounting standards alone is not sufficient for changing firms' actual reporting practices. Especially, enforcement seems to play a crucial role by shaping firms' incentives to produce high quality and comparable financial statements (Christensen *et al.*, 2012; Leuz, 2010; Preiato *et al.*, 2013).

However, for several reasons so far it was not possible to consistently answer the question whether IFRS adoption's objective of higher accounting quality and comparability has been reached. First, only few studies analyze first-order effects of IFRS adoption in a multi-country setting. Second, it is difficult to align the partially contradicting findings of these studies due to fundamental differences in their research designs. Third, research on accounting comparability has been inhibited because for a long time no sound measure of accounting comparability has been available. Finally, to the best of our knowledge to date no comprehensive study examining the effects of IFRS on both accounting quality and comparability as well as the importance of enforcement regulation has been realized. Still, such a combined analysis seems worthwhile because finding an increase in accounting comparability across countries is not necessarily desirable if it is accompanied by a decrease in accounting quality. Therefore, the purpose of this thesis is to analyze the effects of IFRS adoption and enforcement on accounting quality and comparability.

Our sample consists of public firms from 24 countries which adopted IFRS mandatorily before 2009. In order to analyze the influence of enforcement regulation on accounting quality and comparability, we cluster our sample countries into three groups accounting for differences in their institutional enforcement settings. Especially, we differentiate between countries that implemented substantive enforcement reforms concurrent to IFRS adoption from countries that did not change their enforcement systems. To test whether certain effects of IFRS adoption on accounting quality and comparability are of transitory or permanent nature, the sample period of our thesis is split into three subperiods: a four-year pre-adoption period until the date of mandatory IFRS adoption in the EU, a four-year adoption period immediately following IFRS adoption and a four-year post-adoption period thereafter. Thus, our thesis covers an overall time period of twelve years (2001-2012).

In line with prior research, we define high accounting quality as reducing managers' opportunities of earnings management and thus providing decision-useful information to the users of financial statements (Samarasekera *et al.*, 2012). Following Godfrey *et al.* (2010), we define comparable financial statements as being produced consistently, and therefore, allowing users to compare the statements of different firms at one point in time and over time. We perform the tests for accounting quality and comparability independently. In the field of accounting quality, we use a broad set of different metrics which have been validated by prior research in order to test for earnings smoothing, managing towards earnings targets, timely loss recognition, and value relevance. To analyze accounting comparability, we use a relatively new empirical construct which has been developed by De Franco *et al.* (2011).

In all three periods, we find differences both in accounting quality and comparability between our three institutional clusters indicating that enforcement does play a role for the quality and comparability of financial statements. In addition, we find a general drop in accounting quality and comparability after IFRS adoption for our whole sample. This drop does not seem to be of only transitory nature because after the adoption period – which preparers of financial statements might need to adapt to the new accounting standards – accounting quality and comparability as measured by the empirical constructs employed remain lower than before IFRS adoption. Thus, these findings suggest that the positive effects proponents expected from IFRS adoption did not take place on an aggregate level. Furthermore, we find that firms from countries which realized substantial changes in their enforcement settings concurrent to IFRS adoption neither show increases in accounting quality nor in comparability. Hence, the importance of such reforms seems to be limited to second-order effects wherefore the overall role of enforcement seems to be less important than suggested by prior research.

Our thesis contributes to existing literature in several ways. First, by following a comprehensive approach analyzing both accounting quality and comparability we are able to put the findings into a bigger context. Second, prior research only distinguishes between a pre- and post-adoption period, often using only short time periods. By investigating a long-term setting with three periods, our research design permits evaluating whether observed effects are only of temporary or rather permanent nature. Thus, we are able to align partially contradicting findings of prior studies. Furthermore, our thesis is one of the first studies that explicitly analyzes the role of enforcement for accounting comparability in an empirical setting. Simultaneously, it is

one of the few studies based on the empirical construct of accounting comparability developed by De Franco *et al.* (2011), thus adding to the understanding and evaluation of this new measure. Finally, our thesis complements a string of related forthcoming studies on first-order effects of IFRS adoption (e.g. Ahmed *et al.*, 2012; Samarasekera *et al.*, 2012).

The remainder of our thesis is organized as follows. In Section 2, we provide information on the reasons leading to the adoption of IFRS, present the theoretical framework of our thesis as well as the findings of related prior research and discuss the role of a country's institutional setting for reporting practices. In Section 3, we develop our hypotheses. Section 4 introduces our research design. Sections 5 and 6 present our data as well as the results of our tests and are followed by Section 7 which discusses the results. Finally, Section 8 concludes our thesis.

2 Background

In this section, we first explain the motivation of IFRS adoption as well as the concepts of accounting quality and comparability. Then, we discuss potential first- and second-order effects of IFRS followed by the presentation of empirical evidence on such effects and the importance of institutional factors in this context. Finally, we conclude the section stating the implications for our thesis.

2.1 Motivation of IFRS adoption

Financial reporting standards exist because they present an efficient solution to an agency problem which arises from the separation of ownership and control of a firm (Brown and Tarca, 2001; Jensen and Meckling, 1976). In this context, financial reporting standards assist contracting through the standardization of financial information, thereby reducing the need for individual contracts between firms and investors for the provision of information (Preiato *et al.*, 2013). Therefore, the provision of useful information has been a central objective of financial reporting standards for a long time (Gjesdal, 1981).

Still, as the information needs vary across countries, accounting standards do as well. As a consequence, it becomes difficult for users of financial statements to consolidate and compare information of companies from different countries (Prather-Kinsey, 2006). However, the increasing importance of economic cross-border transactions also increases the need of consolidated and comparable information emphasizing the need for harmonization of financial reporting standards.

In order to develop and promote such internationally acceptable financial reporting standards, the International Accounting Standards Committee (IASC) was set up in 1973. In their constitution from 2000, the IASC stated as one of the organization's objectives

> *to develop, in the public interest, a single set of high quality, understandable, enforceable and globally accepted financial reporting standards based upon clearly articulated principles. These standards should require high quality, transparent and comparable information in financial statements and other financial reporting to help*

> *investors, other participants in the world's capital markets and other users of financial information make economic decisions.*
>
> (p. 5, para. 2)

With IFRS, the IASC and its successor organization, the International Accounting Standards Board (IASB), issued a comprehensive set of standards designed to meet these objectives. In the Conceptual Framework, the IASB states as general aim of accounting standards

> *to provide financial information about the reporting entity that is useful to existing and potential investors, lenders and other creditors in making decisions about providing resources to the entity.*
>
> (p. A27, OB2)

The IASB further explains that in order to be useful, financial information needs to be relevant and faithfully represent what it purports to present. Moreover, the usefulness of information increases further if it is comparable, verifiable, timely, and understandable. This notion of faithful representation is in line with scholars' requirements towards high quality accounting standards (Knutson and Napolitano, 1998). Such standards reflect substance rather than legal form, record gains and losses more timely, make earnings more informative, provide more useful balance sheets, and furthermore, limit managers' discretion to smooth earnings, manipulate provisions, and create hidden reserves (e.g. Ball, 2006).

In line with the goal of faithful representation, we define high accounting quality as reducing managers' opportunities of earnings management and thus providing decision-useful information to the users of financial statements (Samarasekera *et al.*, 2012). Furthermore, we define comparable financial statements as being produced consistently, and therefore, allowing users to compare the statements of different firms at one point in time and over time (Godfrey *et al.*, 2010). Comparable financial statements, hence, enable users to identify similarities and differences between entities stemming from economic reasons and not accounting choices (DeFond *et al.*, 2011). Barth *et al.* (2012) specify this notion arguing that if two firms experience similar economic events and report correspondingly similar accounting amounts, the financial reporting of the two firms is comparable.

As worldwide more than 100 countries require or permit IFRS, the adoption of the new financial reporting standards presents one of the major regulatory changes in accounting history (Daske *et al.*, 2008). In the EU for instance, EU Regulation No. 1606/2002 requires listed companies to prepare their consolidated financial statements in accordance with IFRS for fiscal years ending the 31st of December 2005 and later. Not surprisingly, this change was preceded by an extensive debate on the pros and cons of IFRS adoption. One major point of discussion was whether IFRS would indeed lead to increases in accounting quality and comparability as well as positive capital market effects. Theoretically, it can be argued both ways and we present these arguments in the next section.

2.2 Differentiation of first- and second-order effects (of IFRS adoption)

In the following, we present theoretical arguments explaining how IFRS could potentially affect accounting quality and comparability as well as capital markets. In line with prior literature, we

differentiate the effects of IFRS adoption into two types: first-order effects and second-order effects (e.g. Hail *et al.*, 2010). Harmonizing accounting standards by establishing one uniform set of standards has direct effects on the properties of firms' financial statement information, i.e. the quality and comparability of financial statement information. These so-called first-order effects can in turn lead to second-order capital market effects (e.g. changes in market liquidity, firms' cost of capital or the information environment of potential investors) which are the ultimately desired effects of converging accounting standards. *Figure 1* summarizes this categorization of the effects of IFRS adoption.

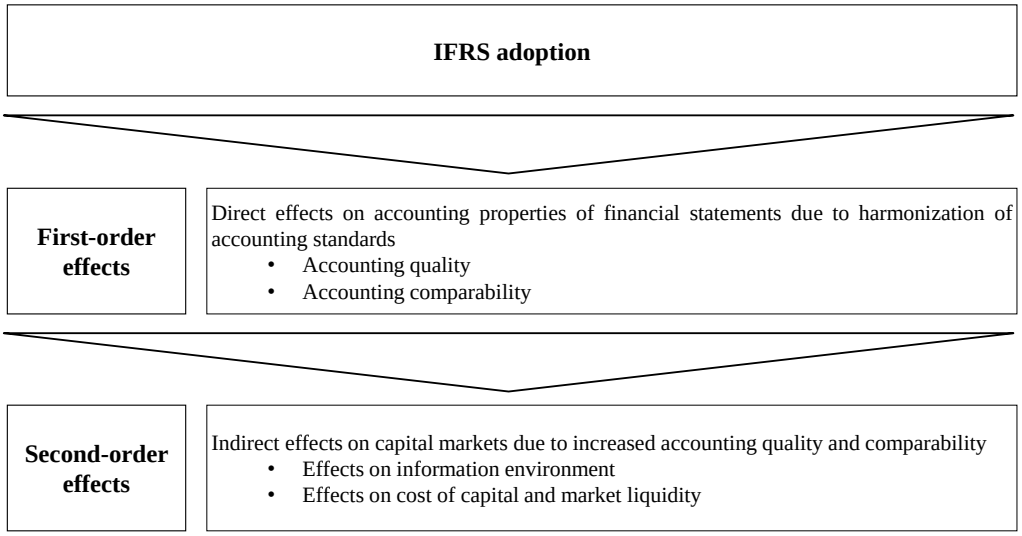

Figure 1: Categorization of the effects of IFRS adoption

Ball (2006) explains that if IFRS are of higher quality than the local generally accepted accounting principles (GAAP), accounting quality (first-order effect) should increase and lead to better-informed valuations in equity-markets, thereby lowering investors' risk (second-order effect). Furthermore, he states that through an increased comparability (first-order effect) IFRS should require less adjustments by analysts in order to make information comparable. Thus, investors' cost of processing information should be reduced and most likely increase stock market efficiency (second-order effects). Moreover, an increase in accounting quality (first-order effect) should reduce information asymmetries between investors which lead to adverse selection as uninformed investors are less willing to trade with better-informed investors and demand a compensation for doing so. Reducing information asymmetries should mitigate adverse-selection and thereby lower firms' cost of capital (second-order effect). Ball (2006) further argues that increases in accounting quality and comparability (first-order effects) lead to an increased scrutiny of management improving corporate decision-making and thereby firm performance.

Although opponents of IFRS adoption do not question the connection between first- and second-order effects, they argue that positive capital market effects might be doubtful as it is not clear whether IFRS present standards of higher quality compared to local GAAP. For example, some authors argue that IFRS as principle-based accounting standards inherently offer certain discretion which will be used differently across firms and countries so that no increase in terms of accounting quality and comparability can be expected (Ormrod and Taylor, 2004). Even if IFRS would decrease accounting discretion, Barth *et al.* (2008) state that the new

standards might be of lower quality if they limit managerial discretion in terms of accounting choices and thus eliminate firms' possibility to report measurements better reflecting firms' actual economic position and performance.

To sum up, there are clear theoretical arguments why improved (reduced) accounting quality and comparability lead to positive (negative) capital market effects. As pointed out, IFRS was introduced with the objective of improving accounting quality and comparability. Therefore, in the next sections we analyze empirical evidence on the effects of IFRS adoption, starting with the second-order effects.

2.3 Empirical evidence on second-order effects

Following IFRS adoption in many countries, academic research started analyzing whether or not IFRS led to the desired effects. As the effects on capital markets are better documented, we first present empirical evidence on second-order effects of IFRS adoption. In general, prior research on second-order capital market effects can be categorized along three dimensions: First, whether IFRS adoption was mandatory or a voluntary choice of firms; second, the sample period and third, one can differentiate between effects on firms' cost of capital and market liquidity on the one hand and effects on the broader information environment on the other hand. In the following, these dimensions are used to structure the existing research.

2.3.1 Empirical evidence on voluntary IFRS adoption

Leuz and Verrecchia (2000) investigate the effects of higher disclosure levels on firms' cost of capital. They argue that opaque reporting practices increase risks for investors and therefore reduce the stock demand, thus increasing bid-ask spreads and decreasing stock turnover. These firms hence have to grant a higher discount to investors, i.e. have higher cost of capital. Using a sample of voluntary IFRS adopters from Germany where local GAAP required lower disclosure levels than IFRS, they indeed find lower bid-ask spreads and higher stock turnover for IFRS adopters. This implies a decreased cost of capital due to IFRS adoption. Daske *et al.* (2011) examine the effects of voluntary adoption on firms' cost of capital and market liquidity. They find little, if not contrary, evidence that – on average – these measures improved. However, they report substantial differences between firms and show that firms' reporting incentives thereby play a major role. In line with Burgstahler *et al.* (2006), reporting incentives are incentives to report informative earnings. Thus, companies can either just adopt the IFRS 'label' or truly commit to a reporting strategy of increased transparency. Only for such 'serious' adopters with high incentives to report informative earnings, Daske *et al.* (2011) could observe positive capital market effects.

Summing up, the prior empirical evidence on voluntary IFRS adoption shows inconsistent capital market effects. Only firms with strong reporting incentives seem to benefit from IFRS adoption, indicating that accounting standards are not the only factor to consider in this context. Moreover, concerns about the validity of these results have been raised (e.g. Soderstrom and Sun, 2007), as the research design of these studies might not control sufficiently for self-selection bias. In other words, it could be that companies adopting IFRS voluntarily did this as a consequence of or concurrent with a general change towards more transparency. Moreover,

the study of Leuz and Verrecchia (2000) might be biased as their sample includes companies listed at the German Neuer Markt which were characterized by spurious business models and reporting practices.

2.3.2 Empirical evidence on mandatory IFRS adoption

2.3.2.1 Effects of IFRS on market liquidity and firms' cost of capital

Daske *et al.* (2008) analyze the effect of IFRS adoption on market liquidity, cost of capital and Tobin's q for a broad sample of mandatory and voluntary adopters from 26 countries between 2001 and 2005. They find an increase in both market liquidity and cost of capital. However, they trace the latter back to anticipating effects and can show a decrease in cost of capital one year before mandatory adoption. Another reason for the increase in cost of capital in 2005 might arise from transitory effects caused by the application of IFRS 1 which made estimates of cash flow and earnings figures temporarily more difficult. Interestingly, Daske *et al.* (2008) show that the capital market effects are stronger for voluntary adopters, also around the time of mandatory adoption, even though these companies had already switched to IFRS before. One reason for this might be concurrent changes in governance, auditing, and enforcement regimes which were implemented in some countries at the same time and which affected both mandatory and voluntary adopters (the concept of complementary factors will gain relevance later in this thesis). Besides, the findings are less pronounced for countries with smaller differences between IFRS and local GAAP as well as countries that started convergence projects before IFRS adoption. Following this direction, the authors show that the capital market effects can only be observed in countries with strong enforcement regimes where the institutional environment provides firms with strong incentives to be transparent. This implies that the capital market effects might primarily not be due to the change of financial reporting standards but rather be a joint effect of this change and complementing changes in the institutional environment. Nonetheless, one limitation of this study is the focus on the pre-adoption phase, so that it becomes not clear whether the results are of temporary or permanent nature.

In another study, Li (2010) reports a significant decrease in the cost of capital for mandatory IFRS adopters. Her sample, however, only includes firms from the EU. While one could argue that this increases the homogeneity of the regulatory background to some extent (thus reducing the influence of other institutional differences on the analysis), the results might be influenced by the relatively high average level of governance quality in the EU. The author finds that the effects can only be observed in countries with strong legal enforcement regimes. In contrast to Daske *et al.* (2008), Li (2010) finds that differences between voluntary and mandatory adopters in cost of capital smooth out after mandatory adoption, also when controlling for transitory effects. However, the sample period from1995 to 2006 might end too soon after the mandatory adoption of IFRS to be able to capture its long-term effects.

DeFond *et al.* (2011) examine the effect of IFRS adoption on cross-border investments by foreign mutual funds. Based on data from 14 EU countries, they investigate the time period from 2003 to 2007 and argue that improved comparability facilitates foreign investment. They find that foreign mutual fund ownership increases when IFRS is faithfully and credibly applied (i.e. no mere 'label' adoption). Beneish *et al.* (2012) also examine the effect of mandatory IFRS adoption on foreign investment differentiating between equity and debt markets and using a

large country sample with data from 2003 to 2007. They find that IFRS adoption affects debt financing decisions to a significantly larger extent than equity financing decisions. Interestingly, only increases in foreign equity investments are limited to countries with high governance quality whereas foreign debt investments are not. This might be due to the fact that debt investors can offset weak governance quality and investor protection by adequate debt covenants. Furthermore, the authors try to distinguish whether these effects stem from increases in accounting quality or comparability. As the authors find an increase of investments from IFRS non-adopting countries and argue that an increase in accounting comparability would have primarily led to an increase of investments from other IFRS adopting countries, they conclude that the observed increase in foreign investment goes mainly back to improvements of accounting quality. However, this method of detecting primary causation seems to be questionable as there is no direct test of the effects.

2.3.2.2 Effects of IFRS on the broader information environment of firms

Armstrong *et al.* (2010) analyze the market reaction to historic events that made the adoption of IFRS in the EU more likely such as evaluations of IASB proposals by EU authorities. Using market return data, they find that the market reacted positively to such events in the pre-adoption phase, indicating anticipated benefits from IFRS adoption. The reactions were especially positive for firms with lower levels of pre-adoption information quality, and hence, higher information asymmetry. Thus, the market expected a better information environment due to uniform accounting standards. Only for firms from code law countries, they documented negative reactions which might have occurred due to market concerns on the compatibility of the enforcement systems in these countries with the new standards. The results are, however, based on the assumption that these changes in probability of IFRS adoption are clearly restricted to the identified events and no information leakage of the policy considerations of EU authorities occurred.

Byard *et al.* (2011) analyze the effect of mandatory IFRS adoption on the information environment of financial analysts, using absolute forecast errors, forecast dispersion, and analyst following as proxies for the quality of the information environment. They find that mandatory adoption on average did not lead to any improvements. Only in countries with strong legal enforcement and local GAAP with substantial differences to IFRS they documented improvements of the information environment. Hence, this study also highlights that the mere change of accounting standards might not be sufficient to alter actual reporting practices, unless there are appropriate enforcement regimes and reporting incentives in place. The findings of this study, however, are only based on the two-year period after the date of mandatory adoption, and furthermore, might be limited by shortcomings of the control group (voluntary adopters).

In another study on analyst following and forecast accuracy, Bae *et al.* (2008) examine whether GAAP differences influence financial analysts' situation. They find a negative correlation between both GAAP differences and analyst following as well as forecast accuracy. This implies that different accounting standards impose costs to foreign financial analysts so that they might decide not to follow a foreign company. Consequently, the harmonization of accounting standards created benefits for financial analysts, thus contributing to a better information environment. One limitation of the authors' approach is that they measure

differences between accounting standards and not actual reporting practices. In other words, voluntary disclosures of companies – though not mandated by an accounting regime – might facilitate analyst following and forecast accuracy but are not included into their study.

Hence, on balance, evidence on mandatory IFRS adoption shows positive second-order effects. However, in almost all studies these benefits are restricted to countries with strong enforcement regimes or high governance quality, and consequently high reporting incentives. This underlines that accounting standards alone are not sufficient to understand the consequences for actual reporting practices. Rather, other factors than accounting standards like the institutional setting might explain the observed effects. This notion is strengthened by the evidence of voluntary IFRS adoption which – though less consistent – hints at the same direction. Moreover, relatively little was known about IFRS' first-order effects as channels of the observed capital market effects. This motivated additional research on the first-order effects of IFRS adoption on accounting quality and comparability, especially taking into consideration countries' institutional setting. Therefore, in the next part we describe how the institutional setting influences accounting quality and comparability (e.g. by shaping firms' reporting incentives) and provide an overview of prior empirical evidence on these first-order effects of IFRS adoption.

2.4 Institutional setting and the empirical evidence on first-order effects

2.4.1 Definition of the institutional setting

Before analyzing the influence of institutional factors on reporting practices, one needs to gain a clear understanding of what is referred to when speaking of a country's institutional setting. To begin with, Olson (1971) defines institutions as mechanisms that facilitate efficient exchanges and interactions between economic players. There is a wide range of definitions, varying in the scope of what is considered an institution (North, 1990; Scott, 2001; Williamson, 1975). However, they generally all agree in that institutions become necessary in order to guarantee efficient exchanges if transaction costs exist. In a world without transaction costs, there would be no need for institutions. But in reality, transaction costs often foster opportunistic behavior which inhibits trade. Institutions then are needed to prevent opportunism and support economic interactions (Coase, 1960; Wysocki, 2011). According to this definition, accounting presents an institution. Taking a transaction cost perspective, accounting facilitates trade between contracting parties by reducing information and coordination costs and improving enforcement of property rights (Barzel, 1982; Watts and Zimmerman, 1986). All institutions of a country form the country's institutional framework which Leuz (2010) defines as follows:

> *A country's institutional infrastructure (or framework) comprises public and private rules, conventions and organizations that shape economic behavior. This includes the legal system, banking system, taxation system, capital markets, regulatory and enforcement agencies, industry associations, standard setting bodies, etc.*
>
> (p. 236)

With this conceptual notion of institutional settings in mind, we can now proceed to investigate how institutional factors influence reporting practices, both in terms of accounting quality and accounting comparability.

2.4.2 Influence of the institutional setting on accounting quality

2.4.2.1 Accounting quality framework

Soderstrom and Sun (2007) develop a schematic framework which depicts the factors influencing accounting quality (*Figure 2*). According to this framework, accounting quality is not only influenced by the ruling accounting standards but also by the legal and political system as well as a number of reporting incentives. We use this framework to provide an overview of these factors and to structure the existing empirical research from this field.

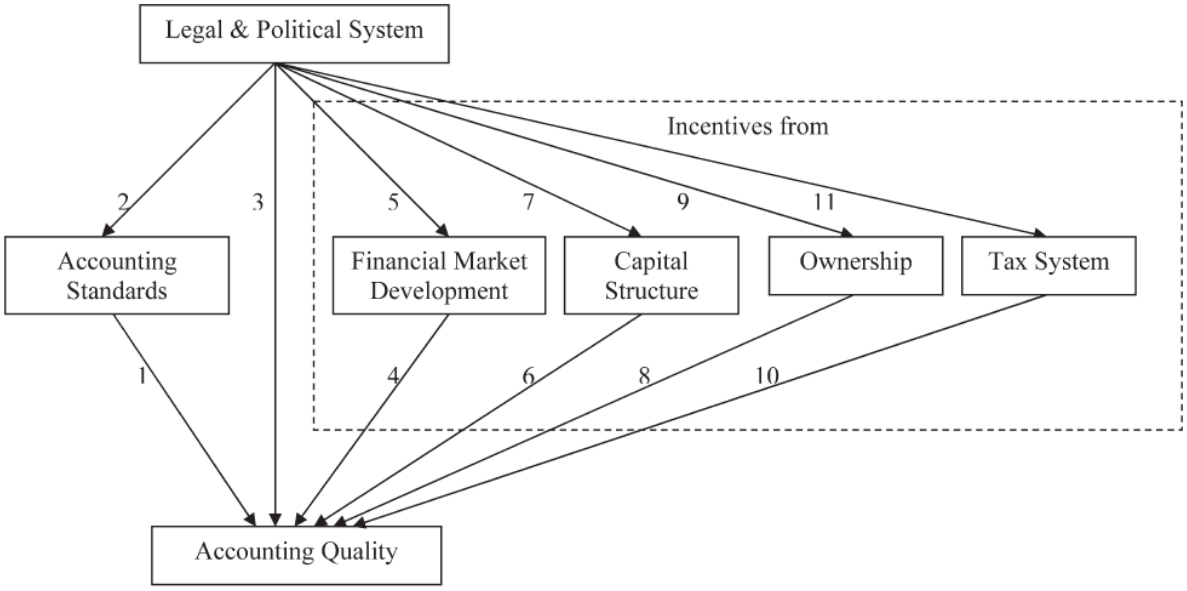

Figure 2: Determinants of accounting quality (Soderstrom and Sun, 2007)

First, the quality of accounting is influenced by the underlying accounting standards (arrow 1 in *Figure 2*). As discussed above, an increase in accounting quality was one of the major objectives of the IFRS adoption across the world. Thereby, financial reporting under high quality standards is value-relevant and reliable. *Figure 2* furthermore shows that the legal and political system influences accounting quality in several ways: to begin with it takes indirect influence through accounting standards (arrow 2 in *Figure 2*). Thereby, the standard setting is usually a political process where different users of financial reports such as banks, shareholders, tax authorities, labor unions etc. are involved. Moreover, the characteristics of the legal system have further implications for accounting standards as witnessed by differences between common and code law countries. While in common law countries the separation of the executive and judicial system leads to accounting standards focusing on information demands from investors, in code law countries accounting standards are part of the commercial law and thus shaped by government priorities (Ball *et al.*, 2003).

As shown by arrow 3 (*Figure 2*), the legal and political system also exerts a direct influence on accounting quality through the legal enforcement of accounting standards. Enforcement aims at fostering compliance with accounting standards in order to assert property rights of

individuals as well as to support broader objectives of market confidence and financial system stability (Carvajal and Elliott, 2009). To achieve this goal, enforcement sanctions misrepresentations and errors of financial statements, thus also deterring future wrongdoing. There are several levels of enforcement which can be differentiated. Thus, enforcement comprises independent enforcement bodies, self-enforcement by preparers and through external auditors as well as the enforcement role of litigations and public sanctions through the media (Fédération des Experts Comptables Européens, 2001). In line with Preiato *et al.* (2013), we focus on enforcement by independent bodies and their monitoring, reviewing and sanctioning activities. Interestingly, in the case of IFRS, there is incongruence between standard setter and enforcement authority. While the IASB issues the financial reporting standards, it does not have enforcement power. Hence, as legal enforcement seems to be a factor influencing accounting quality but which is subject to national regulation, accounting quality is expected to vary across countries although these countries might apply the same accounting standards. This difference in enforcement across countries could be exacerbated through the principle-based nature of IFRS which leaves considerable discretion to the preparers. Enforcement bodies play an important role in shaping this discretion and if this is done differently across countries, differences in accounting quality will increase.

In addition to the above mentioned factors, Soderstrom and Sun (2007) claim that accounting quality is affected by a set of financial reporting incentives. They explain that these incentives arise from the supply and demand for information. As Ball (2002) states:

> *all parties contracting or contemplating contracting with the firm demand information about the firm's ability to meet its contractual obligations. Firms therefore agree to incur the costs of supplying information, and in return they receive better terms of trade from factor owners and customers.*
>
> (p. 131)

Thus, financial reporting presents an equilibrium outcome of the cost of disclosure and the benefits of meeting the information needs of (potentially) contracting parties. Soderstrom and Sun (2007) present the following four financial reporting incentives to be likely to affect accounting quality. First, the development of financial markets plays a role (arrow 4 in *Figure* 2). Thereby, market participants demand information in order to reduce information asymmetry. As explained above, this reduction of information asymmetry is essential in order to reduce the costs of capital for firms. At the same time, the development of financial markets is influenced by the legal and political system which provides the legal frame for investors and thereby determines the degree of investor protection (arrow 5 in *Figure 2*). Second, a firm's capital structure and its corresponding financing requirements influence accounting quality (arrow 6 in *Figure 2*). For instance, firms which are largely financed by bank loans may have a lower incentive to disclose information as public equity markets and the cost of capital at these markets might not affect them. On the other hand, banks use different information channels, and consequently, do not rely on public information disclosure. Again, the legal and political system comes indirectly into play as it influences the capital structure of firms, e.g. through the degree of investor/creditor protection guaranteed by law (arrow 7 in *Figure 2*). Third, ownership is another reporting incentive (arrow 8 in *Figure 2*). Similar to the capital structure,

certain ownership characteristics such as a high degree of ownership concentration might lower a firm's interest (and need) to publicly disclose information. Analogous to the other incentives, the legal and political system can influence ownership as well (arrow 9 in *Figure 2*). Countries with a high degree of investor protection are more likely to show dispersed equity ownership than countries where investors might fear not being able to claim their interest. Fourth, the tax system is an important incentive affecting accounting quality (arrow 10 in *Figure 2*). A strong linkage between tax laws and accounting standards is likely to reduce accounting quality as tax laws are subject to political interests. Moreover, a higher tax rate is likely to result in a higher incentive to report lower earnings, i.e. increasing firms' incentive to manipulate earnings. As arrow 11 (*Figure 2*) shows, the tax system is created by the legal and political system.

2.4.2.2 Empirical evidence

Analogous to the findings of Soderstrom and Sun (2007), Ball *et al.* (2003) state that in their opinion the strong focus of the academic and professional accounting literature on accounting standards, their regional differences, and the efforts of harmonization is confusing as it ignores the fact that the financial reporting practices are – independent of the applied accounting standards – always also subject to reporting incentives of the preparers, i.e. managers and auditors. In the authors' view, these incentives are shaped by economic and political factors. Economic factors go back to the demand for high-quality financial reporting which is determined by a set of aspects such as the volume of public debt and equity markets (i.e. financial market development; arrow 4 in *Figure 2*). On the other hand, political factors include the government's role in shaping and enforcing reporting regulation (arrows 2 and 3 in *Figure 2*) as well as their influence on the tax system (arrow 11 in *Figure 2*). The study focuses on firms from Hong Kong, Malaysia, Singapore, and Thailand because their jurisdictions are based on the common law system (arrow 2 in *Figure 2*) and their accounting standards are generally accepted as high-quality. Nevertheless, they find that in these countries[1] the general quality of financial reporting is low and shows similarities to the characteristics of code law countries. The authors trace this back to the interplay between accounting standards and preparers' reporting incentives and interpret this finding in the way that reporting quality is in the end determined by the economic and political factors shaping preparers' reporting incentives rather than by accounting standards. Hence, the reporting incentives within the dotted line in *Figure 2* dominate the left part of the figure. Moreover, Ball *et al.* (2003) state that differences in reporting incentives across countries inhibit the extent to which a higher international comparability of financial reports can be solely achieved by a harmonization of accounting standards. However, their results need to be interpreted carefully as they are based on a sample of firms from only four countries and as the applied definitions of common law versus code law are rather broad.

Burgstahler *et al.* (2006) examine the role institutional factors and capital market forces play in influencing firms' reporting incentives. As the accounting regulation in the EU before 2005 was based on a firm's legal rather than its listing status, public and private firms were subject to largely the same reporting standards. Nevertheless, as both types of firms experience different capital market forces, analyzing a sample containing both types of firms provides the

[1] For simplicity reasons the study treats Hong Kong as a country although region might have been the more appropriate term.

opportunity to identify the role of incentives and information needs created by public equity markets (arrow 6 in *Figure 2*). In a sample containing companies from 13 EU countries, they find substantial differences in earnings informativeness between private and public companies which can be explained by differences in reporting incentives. More specific, testing for a variety of earnings management practices such as earnings smoothing or accruals manipulation, they find that in countries with developed equity markets private companies are more likely to manage their earnings, i.e. show a lower earnings quality. The authors explain this by a higher incentive for public firms to report informative earnings as this can reduce their cost of capital. This finding contradicts the common assumption that public equity markets increase firms' incentives to manage earnings and shows that the capital structure is a determinant of accounting quality. Moreover, due to the fact that private companies usually are owned by a smaller number of shareholders, the study's findings also support the importance of ownership on accounting quality (arrow 8 in *Figure 2*). As this smaller group of owners usually has privileged access to private information, the importance of public information, and thereby accounting quality, decreases. In addition, the authors document that earnings management is more prevalent in countries with weaker legal systems and enforcement (arrow 3 in *Figure 2*) and in countries with higher tax rates (arrow 10 *Figure 2*).

Leuz *et al.* (2003) find evidence that in countries with a strong investor protection regime earnings management practices tend to decrease. The authors assume that due to conflicting interests between firm insiders and outsiders incentives to manage earnings arise. Thereby, firm insiders such as controlling shareholders or managers could use private information in order to benefit themselves at the expense of other stakeholders. As investor protection limits the insiders' possibility to gain such information, the authors hypothesize that countries having a strong protection of investors will have lower levels of earnings management. Examining a sample of firms from 31 countries between 1990 and 1999, the authors use a set of measures such as earnings smoothing and loss avoidance to detect earnings management. In order to capture the effects from differences and similarities in countries' institutional setting on earnings management, they use variables such as investor protection, ownership concentration, and stock market development to cluster the 31 sample countries into three groups. The first group consists of countries characterized by strong investor rights, dispersed ownership, large stock markets, and strong legal enforcement and contains countries like e.g. the UK or the United States of America. Due to its characteristics the authors label this group *outsider economies*. The other two groups are correspondingly labeled *insider economies* and comprise countries with weaker investor rights, a higher concentration of ownership, and smaller stock markets. Other authors use the labels *market-based* versus *bank-based financial systems* in this respect (Antoniou *et al.*, 2008; Demirgüç-Kunt and Levine, 1999; Ergungor, 2004). The second cluster differs from the third one showing significant stronger legal enforcement. For instance, while Sweden and Germany are part of the second cluster, Greece and Italy are part of the third one. Indeed, Leuz *et al.* (2003) find that firms from countries with strong investor rights, dispersed ownership, developed equity markets and strong legal enforcement are less likely to manage earnings. Hence, they conclude that investor protection is a significant factor for international differences in earnings management, and consequently, accounting quality (arrow 4 in *Figure 2*). However, the results of this study are limited as it is not clear to which extent other institutional factors might account for the explanatory power of the investor

protection variable. Moreover as the study is based on data long before the introduction of IFRS, the 31 sample countries show large differences in accounting standards which cannot be controlled for.

Barth *et al.* (2008) investigate whether the application of International Accounting Standards (IAS) results in a decrease of earnings management. Therefore, they analyze a sample consisting of firms from 21 countries which adopted IAS between 1994 and 2003. They use a number of metrics in order to test for earnings smoothing, accrual manipulation, loss avoidance, timely loss recognition, and value relevance. Interestingly, the authors match their sample firms with firms from the same country and with similar size characteristics but applying different financial reporting standards in order to control for effects not attributable to differences of the financial reporting standards. They find that firms applying IAS are less prone to earnings management and show more timely loss recognition and a higher value relevance. Although the authors attribute the effect on accounting quality to the interaction of different factors such as accounting standards, their interpretation, enforcement, and litigation (arrows 1, 2, 3, 4 in *Figure 2*), their research design is not able to distinguish these effects, and therefore, is limited in explaining the role of different institutional factors for accounting quality. Although the authors include some controls a further caveat stems from the fact that the sample consists largely of voluntary adopters wherefore self-selection bias might be a problem.

Following a similar approach as Barth *et al.* (2008), in a forthcoming study Ahmed *et al.* (2012) also investigate the effects of IFRS adoption on accounting quality. Thereby, they use a sample consisting of 1,600 firms from 20 countries for a period of two years but – in contrast to Barth el al. (2008) – they focus only on firms that adopted IFRS mandatorily. In order to control for economic effects independent to the financial reporting standards, the authors match their sample firms to benchmark firms based on strength of legal enforcement, industry, size, book-to-market, and accounting performance. Their variables test for earnings smoothing, reporting aggressiveness, and earnings management in order to meet or beat targets. Contrasting to Barth *et al.* (2008), Ahmed *et al.* (2012) find that firms adopting IFRS show significant increases in earnings smoothing and accrual manipulation while the timeliness of loss recognition decreases significantly. Surprisingly, their findings hold especially true for firms from countries with strong enforcement. The authors explain this finding by the general difficulty of enforcing principle-based standards emphasizing the importance of implementation guidance. Nevertheless, their results do not come without caveats. Most importantly, their study only analyzes the two-year period after the mandatory adoption of IFRS in the EU in 2005. Therefore, the question arises if the observed effects do persist into the future or if they rather are of transitory nature. The decrease in accounting quality might stem from a decrease in the compatibility of accounting standards and institutions. However, it could be that this effect is compensated over time by learning effects of the preparers, better implementation guidance or changes in a country's enforcement regulation. Furthermore, the authors' findings are limited by their sample composition where weak enforcement countries represent less than a fifth of the overall observations and are therefore likely to be underrepresented.

Samarasekera *et al.* (2012) examine the role of enforcement for accounting quality (arrow 3 in *Figure 2*). They use a sample consisting of 495 firms from the UK for the period from 2000 to

2009. Furthermore, the sample includes some cross-listed firms as they are subject to more than one enforcement body and therefore allows the authors to test the hypothesis that a higher scrutiny by enforcement bodies leads to a higher accounting quality. In order to proxy for accounting quality, the authors test measures for earnings smoothing, meeting or beating earnings targets, timely loss recognition, and value relevance. Their results show that IFRS adoption leads to a higher value relevance and a lower likelihood of earnings management for meeting or beating earnings targets. However, only for cross-listed companies they can find lower levels of earnings smoothing and a more timely loss recognition supporting their hypothesis and showing that enforcement has a positive effect on accounting quality.

2.4.3 Influence of the institutional setting on accounting comparability

While the majority of prior studies either investigates the effects of IFRS adoption on capital markets or accounting quality, only few studies aim to shed light on its effect on cross-country accounting comparability. One of the reasons for this is that for a long time no empirical measure for financial statement comparability existed (De Franco *et al.*, 2011). For instance, as mentioned earlier DeFond *et al.* (2011) argue that increases in foreign mutual fund ownership can be explained by an increase in accounting comparability. They measure accounting comparability as the number of industry peers using the same accounting standards, provided the latter are credibly implemented. However, other research proved that merely adopting new accounting standards is insufficient in changing actual reporting practices (Daske *et al.*, 2008; Li, 2010). Hence, one might raise concerns about taking the number of peers using the same standards as a measure of accounting comparability. DeFond *et al.* (2011) therefore introduce the concept of implementation credibility and use the earnings quality score from Leuz *et al.* (2003) as a proxy. This measure, though, is rather a measure of accounting quality than comparability so that no clear separation between both first-order effects is achieved.

Finally, De Franco *et al.* (2011) developed an empirical construct of accounting comparability which has gained considerable attention recently. Based on the notion that accounting systems are mappings from economic events to financial statements, they argue that firms having similar mappings have comparable accounting systems. Thereby, firms with comparable accounting systems produce similar financial reports if they experience similar economic events. Their measure does not rely on a comparison of input data of financial statements (such as the choice of accounting methods applied) but rather uses the output of financial statements (e.g. earnings and return figures). In doing so, they use stock returns and earnings as proxies for economic events and output of financial statements. Furthermore, they prove the validity of their empirical construct by showing that the measure is positively related to analyst following and negatively related to analysts' forecast dispersion (which are theoretically linked to greater accounting comparability).

Following De Franco *et al.* (2011), researchers started applying the new measure to directly test for the effects of IFRS adoption on accounting comparability. However, this research is still in its infancy. Cascino and Gassen (2012) analyze the accounting comparability effects of mandatory IFRS adoption based on the methodology proposed by De Franco *et al.* (2011), using a sample of 14 countries that adopted IFRS and 15 benchmark countries for the period of 2001 to 2008. Further, they extend the methodology suggested by De Franco *et al.* (2011) by also

using cash flows instead of stock returns as proxies for economic events. The authors document no clear positive effect on accounting comparability and try to explain these findings by performing a more detailed comparison between IFRS compliance of German and Italian firms. They then argue that the adoption of harmonized accounting standards can only affect accounting comparability if firms comply with these standards. Based on their analysis of firm-, region- and country-level incentives for accounting compliance (as assessed by hand-collected data), they show that the accounting comparability effect of IFRS adoption varies systematically with compliance incentives. Only firms with high compliance incentives experience significant increases in accounting comparability.

Lang *et al.* (2010) similarly investigate the effects of mandatory IFRS adoption on accounting comparability using the approach of De Franco *et al.* (2011). However, they also consider the concept of earnings co-movement as another measure of comparability. Nevertheless, they point out that earnings co-movements capture anything that leads to similar earnings (not only similar underlying economic events). Therefore, similar earnings as documented by this measure do not necessarily indicate comparable financial statements. Despite these concerns about their research design, they find that accounting comparability measured with the approach of De Franco *et al.* (2011) decreases after IFRS adoption compared to a benchmark of non-adopting countries. They argue that local GAAP reflects differences in the local institutional environment and therefore introducing harmonized accounting standards might extinguish differences in accounting choices which might have been legitimate. Still, unlike Cascino and Gassen (2012) they do not follow the idea of a potential impact of institutional differences on accounting comparability.

Brochet *et al.* (2012) indirectly examine the effects of mandatory IFRS adoption on accounting comparability. Using a UK-based sample where local GAAP is quite similar to IFRS, they argue that any capital market effects following IFRS adoption are likely to primarily stem from improved accounting comparability rather than quality. They find that IFRS adoption leads to a decrease of information asymmetries between public and private information (documented as a decrease of insiders' ability to benefit from insider purchases) and argue that this implies an increase of accounting comparability. This is inconsistent with the studies previously presented. Interestingly, they also test for changes in accounting comparability using the approach by De Franco *et al.* (2011) as applied in Lang *et al.* (2010). According to them, the test design of De Franco *et al.* (2011) overall implies an increase in accounting comparability following IFRS adoption. However, as they employ a one-country setting, no inferences on a potential influence of institutional differences can be made.

To conclude, the empirical evidence on the effects of IFRS adoption on accounting comparability is inconsistent, mostly indicating only minor improvements at best. Moreover, it became evident that further research is required to investigate the influence of the institutional setting on accounting comparability. More specific, while Cascino and Gassen (2012) examined the role of reporting incentives in the context of accounting comparability, no study has been conducted so far that analyzes the role of enforcement changes on accounting comparability. In addition, applying the methodology suggested by De Franco *et al.* (2011) to measure comparability effects seems to be promising, but has not been sufficiently explored by prior

research. Thus, measuring accounting comparability effects of IFRS adoption depending on changes in enforcement will be one focus of our thesis.

2.4.4 Institutional complementarities

The discussion of the influence of the institutional setting on accounting quality and comparability illustrates that the individual elements of the institutional infrastructure are interdependent. For instance, Leuz (2010) points out that reporting standards cannot be considered without thinking about how to enforce them. If a standard offers a lot of discretion, then enforcement can become practically impossible. Thus, the author explains that there are considerable interdependencies between the regulatory elements of an institutional setting. Using the notion of institutional complementarities, Leuz (2010) argues that in well-functioning institutional settings, financial reporting and other institutional elements are designed to tie together (see also Hail *et al.*, 2010). This is in line with Wysocki (2011) who states that accounting emerges endogenously in an economy. From this notion of complementarities, it follows that changes in accounting standards cannot be considered in isolation from other elements of the institutional framework. Even if accounting standards *per se* might be of higher quality, institutional incompatibility of changes in these standards with other elements of the institutional setting might prevent improvements in the first-order effects of IFRS adoption. Further, Leuz (2010) finds that there are robust country clusters of institutional differences (see also Leuz *et al.*, 2003) which explain the existence of substantial differences in reporting practices between countries. As the broader institutional infrastructures are sticky over time, the author argues that despite the harmonization of accounting standards differences in reporting practices, and hence in accounting quality and comparability, are likely to persist into the future.

The study of Christensen *et al.* (2012) is an example of how research can use the notion of complementarities to further investigate the influence of the institutional setting on reporting practices. To disentangle the impact of IFRS adoption and concurrent enforcement changes, the authors use data from 35 IFRS adopting countries from 2001 to 2009 in order to analyze changes of market liquidity as a proxy for overall capital market effects. The enforcement reforms in some countries were undertaken to guarantee that enforcement mechanisms were compatible with the new financial reporting standards. This objective of the reforms explicitly acknowledges the existence of interdependencies between accounting standards and other institutional elements. However, while EU Regulation No. 1606/2002 requires countries to adjust their enforcement systems to ensure compliance with IFRS, no detailed guidance on how to do these adjustments is given and considerable discretion on country-level exists. Hence, Christensen *et al.* (2012) find that only six countries implemented substantive changes of their enforcement regimes concurrent to IFRS adoption. Combining secondary data analysis with surveying and interviewing experts in the field of enforcement (e.g. enforcement bodies, auditors etc.), they define substantive changes as a combination of several of the following events: the creation of new enforcement agencies, higher penalties for violations of financial reporting standards, an increase in the resources of the enforcement bodies, and especially the introduction of a proactive review process of financial statements. Thus, the authors find that increases in market liquidity following IFRS adoption are limited to the countries which also changed their enforcement systems simultaneously to IFRS adoption. Their findings especially

show that even in countries with qualitatively high regulatory characteristics, positive effects on market liquidity could not be documented without concurrent changes to enforcement (see also Samarasekera *et al.*, 2012). This is in contrast to prior studies (most notably Daske *et al.*, 2008; Li, 2010) which also analyze IFRS' effects on market liquidity and find an influence of the strength of enforcement. Nevertheless, the studies do not analyze the role of adjustments to enforcement which might be necessary due to the introduction of new financial reporting standards. However, Christensen *et al.* (2012) solely analyze the effect of enforcement changes and concurrent enforcement changes on market liquidity. To better understand the importance of enforcement and its changes in the context of IFRS adoption, a study focusing on the first-order effects of accounting quality and comparability is required.

2.5 Conclusion and implications for our thesis

Summing up, the evidence of the effects of IFRS adoption is mixed. While most prior research finds positive second-order capital market effects following IFRS adoption (though often restricted to subsamples of countries, e.g. with concurrent enforcement changes), there is inconsistent evidence of the first-order effects on accounting quality and comparability. Some studies even imply decreases in accounting quality and comparability. One common finding of many studies is that the mere introduction of harmonized accounting standards is not enough to create positive first- or second-order effects. Rather, it is shown that the institutional setting of countries plays an important role for these effects to take place. Hence, accounting standards are just one piece in a mosaic of mutually interdependent institutional factors which influence reporting practices. Cascino and Gassen (2012) show that the influence of institutional factors is not limited to accounting quality but also valid for comparability. Thus, legal and political systems as well as reporting incentives determine how comparable reporting practices are. For this reason, one could illustrate the determinants of accounting comparability analogously to the framework of Soderstrom and Sun (2007) of the determinants of accounting quality (*Figure 2)*.

The review of prior research also reveals that current research on the influence of the institutional setting on reporting practices in some parts is still in its infancy, particularly in the area of accounting comparability. In this context, especially enforcement seems to be a key factor by shaping firms' reporting incentives. Therefore, understanding the relative roles of enforcement and accounting standards on first-order effects is of utmost importance for accounting research. Hence, the research question of our thesis is how IFRS adoption and enforcement affect accounting quality and comparability as sources of the observed positive capital market effects.

Answering this research question aims at filling three gaps in existing literature: First, previous studies mainly focus on single measures of either accounting quality or comparability. Nonetheless, it is important to test both first-order effects in a comprehensive setting since an increase in comparability might not be desirable if it is accompanied by a reduction of accounting quality. Therefore, a complementary investigation of both effects is vital. Second, prior studies have mainly focused on a short time period around IFRS adoption. Only limited evidence exists covering the development over the full seven-year period since IFRS were adopted mandatorily in the EU. Especially, Ahmed *et al.* (2012) show interesting findings

contrary to the results of other studies but only focus on a two-year period after IFRS adoption. Analyzing longer sample periods might answer the question whether there is a convergence of reporting practices after a necessary adoption phase or rather differences in accounting quality and comparability between certain country clusters continue to exist as predicted by Leuz (2010). Finally, while there have been studies on the influence of enforcement on accounting quality, no such study exists for accounting comparability. In this context, the methodology developed by De Franco *et al.* (2011) provides a promising empirical construct for testing accounting comparability.

3 Hypotheses development

Based on the findings of prior research, we develop six hypotheses about the first-order effects of IFRS adoption. Thereby, we take into account that prior research indicates that the mere introduction of harmonized accounting standards is not enough to create positive first-order effects. Instead, the institutional setting, more specific the enforcement regulation and its effect on reporting incentives, seem to be of particular importance wherefore we must take these institutional differences into consideration. Therefore, we differentiate the firms of our sample and their respective institutional environments on a country level in different groups. Drawing on Christensen *et al.* (2012), we know that some countries realized enforcement reforms in order to adjust their enforcement mechanisms to the new set of standards provided through the introduction of IFRS. Thus, firms from these *enforcement reform countries* (ERC) form a group. Further, we know from prior research (e.g. La Porta *et al.*, 2006; Leuz, 2010) that countries' enforcement strength can generally be differentiated into stronger and weaker settings wherefore we differentiate two further groups: firms from *strong enforcement countries* (SEC) and firms from *weak enforcement countries* (WEC).

3.1 Hypotheses concerning accounting quality

Soderstrom and Sun (2007) explain the determinants of accounting quality and show how enforcement structures directly influence firms' reporting incentives. This argumentation is further strengthened by Christensen *et al.* (2012). Hence, for firms from ERC, the changes in enforcement or the bundled effect of changing accounting standards and conducting concurrent enforcement reforms are predicted to lead to higher reporting incentives which result in higher accounting quality. Accordingly, we test the following hypothesis:

> **H1:** *Firms from ERC will show higher levels of accounting quality after the adoption of IFRS.*

In other words, we predict accounting quality for these countries to increase during an adoption period after the introduction of IFRS and then reach a steady level after this adoption period.

Again referring to the line of reasoning in Soderstrom and Sun (2007), firms from SEC are already used to high quality standards as well as strong and compatible enforcement structures. Therefore, high incentives to report high quality financial statements already exist. For this

reason, we predict that there will be no significant increase in accounting quality in the long run compared to the time before IFRS adoption and test the following hypothesis:

H2: *Firms from SEC will show no significant increase in accounting quality in the long run.*

However, it might be possible that the preparers of financial statements (i.e. firms and auditors) need some time to adjust and adapt their practices to the new standards. There might be an initial general confusion paired with uncertainty on the side of both preparers and investors during the adoption period amplified by the effects of IFRS 1. In this context, studies of Cormier *et al.* (2009) as well as García Osma and Pope (2011) show that the strategic use of first-time adoption adjustments of IFRS 1 created additional transitory effects. For this reason, we focus on a longer time period to enable us to capture potential temporary adoption effects due to learning effects. Nevertheless, we do not develop any hypotheses concerning the temporary effects on accounting quality during an adoption period.

With regard to WEC, although these countries have experienced a change to a new accounting regime, this change is restricted to a single aspect of the institutional framework shaping reporting incentives and continues to be dominated by other institutional factors. As a result, we predict reporting incentives to be unaltered wherefore firms do not significantly change their reporting practices. Hence, we hypothesize:

H3: *Firms from WEC will show no significant increase in accounting quality in the long run.*

Summarizing our hypotheses for accounting quality, we predict a continuing divergence of reporting practices as institutional differences are likely to remain after IFRS adoption. This is in line with Leuz (2010) who argues that as long as there is no convergence of institutional factors there will be no convergence of reporting practices. Next, we will develop similar hypotheses for the development of accounting comparability between the different institutional groups.

3.2 Hypotheses concerning accounting comparability

Regarding the comparability of accounting information of firms from ERC and SEC, IFRS adoption leads to the application of the same financial reporting standards in these countries and the enforcement reforms in ERC are likely to foster a convergence of the institutional framework between the two groups of countries. These enforcement changes also result in a similar interpretation of discretionary choices by firms as the reforms in ERC should lead to a convergence of reporting incentives between the two clusters. For this reason, we predict actual reporting practices to converge and test the following hypothesis:

H4: *Accounting comparability between firms from ERC and SEC will increase after IFRS adoption.*

For firms from WEC, incentives to report high quality financial statements remain low after IFRS adoption and there will be discretion which firms can use opportunistically. Hence, in the

long run, in WEC the discretion will still be used in different ways than in ERC which are converging toward SEC, so that accounting comparability between ERC and WEC predictably will not increase in the long-term. Our hypothesis, therefore, is as follows:

> **H5:** *Accounting comparability between firms from ERC and WEC will show no significant increase in the long run.*

Using the same reasoning, accounting comparability between *SEC* and *WEC* is predicted to remain unaltered in the long run, wherefore we hypothesize:

> **H6:** *Accounting comparability between firms from SEC and WEC will show no significant increase in the long run.*

However, there might also be temporary adoption effects, e.g. because of the accounting choices provided by IFRS 1. However, we do not hypothesize any transitional adoption effects on accounting comparability in the adoption period.

4 Research design

In this section, we first present the sample we use to test our hypotheses. Second, we cluster the firms of our sample into different groups of countries accounting for the differences in the enforcement setting of these countries. Then, we differentiate the sample period into subperiods according to the stage in the process of IFRS adoption. Finally, we present the metrics which we use to test for accounting quality and comparability.

4.1 Sample Selection

As the purpose of our thesis is to analyze the role of countries' enforcement settings in conjunction with the adoption of IFRS for accounting quality and comparability, our sample necessarily consists of firms from countries which adopted IFRS as financial reporting standards. As some of the metrics we apply to test our hypotheses require a minimum time span of four years to deliver reliable results,[2] we exclude firms from countries which adopted IFRS after 2008 because for these firms we would not be able to measure certain effects of the adoption. This reduces the list of potential sample countries to 35 (see *Appendix 1*). However, in order to cluster these countries according to their institutional and enforcement characteristics, we rely on data from prior research which are not available for eleven of these countries wherefore the number of sample countries is reduced to 24 (see *Appendix 2*).

4.2 Clustering of countries

As mentioned before, prior research shows the importance of enforcement in addition to harmonized financial reporting standards for creating positive first-order effects. Moreover, Leuz (2010) finds that countries comprise robust institutional clusters around the world, especially related to enforcement differences. To account for these institutional differences and

[2] For instance, our measure of comparability needs to be based on periods of (at least) four years.

analyze their effects on accounting quality and comparability, we partition our sample into three groups.

Thereby, the first group of countries we distinguish comprises countries which adopted IFRS and realized substantive reforms in their enforcement regulation. In line with Christensen *et al.* (2012), we define substantive enforcement reforms as one or a combination of the following events: the creation of new enforcement agencies, the introduction of a proactive review process of financial statements, higher penalties for violations of financial reporting standards, or an increase in the resources of the enforcement bodies. Christensen *et al.* (2012) find that five countries, namely Finland, Germany, the Netherlands, Norway, and UK, realized such reforms concurrent with the introduction of IFRS.[3] Therefore, these countries form our first cluster which we label *enforcement reform countries* (ERC).

Consequently, the remaining sample countries did not realize such enforcement reforms when introducing IFRS. However, from an institutional perspective these countries cannot be described as being a homogeneous group as there are too big differences in their general institutional frameworks as well as their enforcement mechanisms, and therefore, in their firms' reporting incentives. Hence to identify these differences and perform a further clustering accounting for these differences, similar to Leuz (2010) we use data of institutional characteristics collected by La Porta *et al.* (2006) and Djankov et al. (2008). In total, we use six variables to proxy for a country's strength of financial reporting enforcement (see also *Appendix 3*). The first three variables we draw from La Porta *et al.* (2006) and relate to countries' security regulation laws. More specific, these variables are disclosure requirements, liability standard, and public enforcement. Disclosure requirements is an index comprised of a set of variables related to the level of disclosure in securities offerings such as the obligation to publish a prospectus, the obligation to disclose major shareholders or directors' ownership etc. Liability standard relates to the requirement that the prospectus of a public security offering must include all information required in order to assess the value of the security offered. The variable is an index measuring the strength of liability standards of different groups, namely the issuer of the security, the distributor of the prospectus, and the accountants. The third variable, public enforcement, is an index focusing on different enforcement aspects related to security markets. It is comprised of five subindices such as rule-making power index, investigative powers index, and orders index. The remaining three variables are taken from the study by Djankov *et al.* (2008) and relate to enforcement of anti-self-dealing rights. Self-dealing refers to the expropriation of minority shareholders by corporate insiders (such as managers or controlling shareholders). The first two variables are measures of private enforcement. One refers to the ex-ante private control of self-dealing (i.e. circumstances that deter insiders from self-dealing, such as the amount of immediate, transaction-related disclosure and approval requirements), and the other one to the ex-post control of self-dealing (i.e. the ease of proving wrongdoing for outside shareholders, e.g. access to critical documents). The third variable

[3] Actually, Iceland is found to be a sixth country implementing substantive enforcement changes. However, as data on certain institutional characteristics which is important for the further clustering is not available for Iceland, we do not include it into our analysis.

refers to the public enforcement of self-dealing (i.e. the severity of fines and prison terms for self-dealing).

Thus, in order to assign the sample countries which did not realize enforcement reforms concurrent to the IFRS adoption into a group of *strong enforcement countries* (SEC) or *weak enforcement countries* (WEC), we partition the countries using a k-means cluster analysis based on data on the six institutional variables described above (see also *Appendix 4*), ex-ante specifying two country clusters.[4] *Table 1* shows the resulting allocation of countries to the SEC and WEC cluster from our cluster analysis.

Strong enforcement countries (SEC)	**Weak enforcement countries (WEC)**	
Australia	Austria	Pakistan
Hong Kong	Belgium	Philippines
New Zealand	Denmark	Portugal
Singapore	France	Spain
South Africa	Greece	Sweden
	Israel	Switzerland
	Italy	Turkey

Table 1: Results of k-means cluster analysis

The SEC cluster can be interpreted as a group of countries frequently characterized by large stock markets, low ownership concentration, high investor protection and, most importantly in our analysis, strong legal enforcement. Strong legal enforcement refers to both public and private enforcement as well as factors that facilitate enforcement such as high disclosure levels or the availability of class-action lawsuits. Often, these countries have an Anglo-Saxon legal origin so that local GAAP was similar to IFRS and the enforcement bodies can be expected to already be quite compatible to the new standards. This cluster in many features resembles the conceptual 'outsider economy' as presented by Leuz (2010).

The WEC cluster, on the other hand, can be interpreted as a group of countries frequently characterized by smaller stock markets (i.e. a higher dependence on internal, relationship-based financing), higher ownership concentration, weaker investor protection and lower disclosure levels. Using the terminology of Leuz (2010), these countries often carry traits of 'insider economies' and, in our case, have a continental European legal origin. Legal enforcement, in general, is weaker than in the SEC cluster. Or, in case of countries with strong legal systems, the enforcement structures can be expected to be less compatible with the new standards as the old accounting regimes often were more conservative and less market-oriented.

Our clustering differs from the clusters proposed by Leuz (2010) in two ways. First, Leuz analyzed a broader range of countries, also including non-IFRS-adopting countries which we excluded from our analysis. Second, in the cluster analysis we only build two clusters instead

[4] All statistical calculations of our thesis are carried out in STATA.

of three. *Table 2* shows the three clusters which we will differentiate in our tests of accounting quality and comparability.

Enforcement-reform countries (ERC)	Strong enforcement countries (SEC)	Weak enforcement countries (WEC)	
Finland	Australia	Austria	Pakistan
Germany	Hong Kong	Belgium	Philippines
Netherlands	New Zealand	Denmark	Portugal
Norway	Singapore	France	Spain
United Kingdom	South Africa	Greece	Sweden
		Israel	Switzerland
		Italy	Turkey

Table 2: Country clusters

An alternative procedure of clustering the sample countries would be to include all countries (i.e. also the countries now ex-ante assigned to ERC) into the k-means cluster analysis and then differentiate between two or three clusters with varying strengths of enforcement regimes. However, drawing on Christensen *et al.* (2012) we decided to build a separate cluster for the enforcement reform countries. This is because their findings indicate that capital-market effects following IFRS adoption are restricted to countries implementing concurrent changes in their enforcement regimes. It seemed more insightful to us to investigate whether both the strength of enforcement and changes in enforcement reforms influenced first-order effects of IFRS adoption.[5]

4.3 Time periods

In our research design, we differentiate between three time periods, the pre-adoption period before IFRS adoption (generally 2001-2004), the adoption period (generally 2005-2008) and the post-adoption period (generally 2009-2012). The time spans of the periods are contingent on the metrics of our research design and data availability. For example, the accounting comparability metric leads to better results for four-year periods. We adjust the periods if firms' fiscal years differ from calendar years.[6] Other studies that investigate a pre/post-IFRS design (e.g. Ahmed *et al.*, 2012; Cascino and Gassen, 2012; Lang *et al.*, 2010) only distinguish between a period before and after IFRS adoption. A split into three periods, however, appears reasonable to investigate whether observed effects are of temporary or rather permanent nature. For instance, Ahmed *et al.* (2012) find evidence contrasting the findings from Barth *et al.* (2008) but based only on an analysis of the two-year period following IFRS adoption. Hence, it is

[5] However, if the ERC countries are included into the k-means cluster analysis, the resulting clusters do not differ except for the United Kingdom which then belongs to the SEC cluster. Therefore, it might be that assigning the UK to a specific cluster has an impact on the results. For this reason, we later perform a robustness test leaving the UK out of the sample to see whether or not our inferences are changed.

[6] However, we do not adjust the periods for the few non EU-countries that have different IFRS adoption dates (see *Appendix 1*) because when using relative periods results might be distorted by general economic trends not being correctly captured. Rather, we perform a robustness test by leaving those countries out of the analysis to see whether results are driven by them.

unclear whether the observed effects continue to exist or level off after a necessary adaption phase preparers might need to get used to the new standards. For this reason, we need to split the sample period into three subperiods to see whether or not temporary developments during the adoption phase level off in the last period.

4.4 Accounting quality

In order to test hypotheses H1 to H3 which refer to accounting quality, we use a set of metrics covering different aspects of accounting quality. As all of these three hypotheses refer to the development of accounting quality over time, for each hypothesis the overall development of these metrics will be taken into account. Furthermore, these metrics do not only measure the effect from accounting standards and other institutional variables such as the enforcement regulation but also capture other factors such as the economic environment. In order to mitigate the influence of these effects – which are not attributable to the institutional design we want to analyze – we include control variables which were identified by prior literature as related to the economic environment.

The accounting quality metrics we use can be differentiated into two groups: earnings management and value relevance. For each of these two groups, we use metrics identified and applied by prior research (Ahmed *et al.*, 2012; Barth *et al.*, 2008; Lang *et al.*, 2006; Samarasekera et al., 2012). The reasoning behind this decision is that these metrics were not only validated but that their application further presents a necessary condition in order to be able to reconcile the partially contradicting findings of prior studies. As earnings management is difficult to measure empirically, the approach we take with our research design is necessarily indirect. For all metrics, we first estimate separate regressions for each cluster and time period. In the case of two of the earnings management metrics (the frequency of small positive and large negative net income), our final metric is a coefficient of the respective regressions. Therefore, in these cases we test for significance by simply using a t-test. For the other metrics (the remaining earnings management and the value relevance metrics), we are interested in the differences in the metrics between clusters and periods (e.g. whether firms from ERC show a significantly higher result of the metric than firms from WEC). To test whether these differences are significant, however, it is not possible to simply use a t-test because the empirical distribution of the differences is unknown. Consequently, we first need to generate the empirical distribution of the differences with the help of a percentile bootstrapping approach and can then in the next step use a t-test to calculate whether the metrics differ significantly between clusters and periods.

For that reason, in line with Barth *et al.* (2008) and Ahmed *et al.* (2012) we use a percentile bootstrapping approach for three of the earnings management metrics (the three earnings smoothing metrics) and the value relevance metrics in order to generate the empirical distributions of the metrics. Thereby, we randomly select, with replacement, observations from each cluster-period pair to build representative samples that are as large as our actual sample. Then, we calculate the metrics and the differences in each metric between clusters and time periods. By repeating this procedure 100 times, we obtain the empirical distribution of these differences. Based on them, we use t-tests to investigate whether the differences between

clusters and periods are significant. We consider a difference significant at the p<10% (p<5%, p<1%) level, if a confidence interval bounded by the 5th and 95th (2.5th and 97.5th, 0.5th and 99.5th) percentile of the bootstrap distribution fails to contain zero.

In the following, we present the metrics which we use in our analysis of accounting quality. First, we introduce the five metrics of earnings management and thereafter, we present the three value relevance metrics.[7]

4.4.1 Earnings management

There are different facets of earnings management. For example, firms can manage earnings either by smoothing their earnings figures, managing towards earnings targets or avoiding a timely recognition of losses. Therefore, we use three different metrics capturing earnings smoothing and one for managing towards earnings targets and timely loss recognition, respectively. After computing these measures we combine the results in order to arrive at one conclusion about the level of earnings management done in each cluster and period. However, unlike other research (Leuz *et al.*, 2003), we do not aggregate these metrics into one overall earnings management measure because such an aggregation leads to a loss of important information on a more detailed level (e.g. differences between the five earnings management metrics).

4.4.1.1 Earnings smoothing

The first of the three metrics (metric 1.1, see *Appendix 5*) which we use for earnings smoothing is the variability of net income scaled by total assets. If firms smooth earnings they will show a lower variability of net income. As net income is subject to a variety of other factors, we include a set of control variables. In line with prior research (e.g. Barth *et al.*, 2008; Lang *et al.*, 2006), we operationalize this by first estimating a regression of the change in annual net income (scaled by total assets) on our control variables. Then, we use the residuals from this regression to calculate the measure of variability of net income. Thus, *variability in ΔNI* is the variance of the residuals from a regression of the change in annual net income scaled by total assets on the control variables.[8] Greater variance indicates higher accounting quality.

$$\Delta NI_{it} = \alpha_0 + \alpha_1 SIZE_{it} + \alpha_2 GROWTH_{it} + \alpha_3 EISSUE_{it} + \alpha_4 LEV_{it} + \alpha_5 DISSUE_{it} + \alpha_6 TURN_{it} + \alpha_7 CF_{it} + \alpha_8 AUD_{it} + \alpha_9 ADR_{it} + \alpha_{10} CLOSE_{it} + \varepsilon_{it} \quad (1)$$

where:

SIZE = natural logarithm of market value of equity by end of fiscal year;
GROWTH = annual percentage change in sales;
EISSUE = annual percentage change in book value of equity;
LEV = total liabilities divided by book value of equity at fiscal year-end;

---

[7] An overview of the metrics can be found in *Appendix 5*.

[8] Barth *et al.* (2008) use *NUMEX* – the number of exchanges on which a firm's stock is listed – and *XLIST* – an indicator variable that equals one if the firm is listed on any U.S. stock exchange and Worldscope indicates that the U.S. exchange is not the firm's primary exchange, and zero otherwise. However, as no data on these variables was available in DataStream we use *ADR* – American Depository Receipt – which also represents a form of cross-listing in the US.

DISSUE = annual percentage change in total liabilities;
TURN = annual sales divided by end of fiscal year total assets;
CF = annual net cash flow from operating activities divided by end-of-year total assets;
AUD = an indicator variable that equals one if the firm's auditor is PwC, KPMG, Arthur Andersen, E&Y, or D&T, and zero otherwise;
ADR = an indicator variable that equals one if the firm trades American depository shares, hence a form of cross-listing;
CLOSE = the percentage of closely held shares of the firm as reported by Worldscope.

Our second earnings smoothing metric (metric 1.2, see *Appendix 5*) tries to account for the fact that variance in net income may stem from firm specific factors which also affect a firm's underlying cash flows. Naturally, firms with higher variability of cash flows will also show a higher variability of net income. Therefore, we build the mean ratio of the variability of the change in net income to the variability of the change in operating cash flows. If the variance of net income does not stem from a corresponding variance in operating cash flows but rather from the fact that a firm uses accruals to manage earnings, the variability in net income should be lower than the one in operating cash flows. Similar to the change in net income, change in operating cash flow is also sensitive to a variety of factor unrelated to the institutional framework wherefore we estimate the following regression:

$$\Delta CF_{it} = \alpha_0 + \alpha_1 SIZE_{it} + \alpha_2 GROWTH_{it} + \alpha_3 EISSUE_{it} + \alpha_4 LEV_{it} + \alpha_5 DISSUE_{it} + \alpha_6 TURN_{it} + \alpha_7 CF_{it} + \alpha_8 AUD_{it} + \alpha_9 ADR_{it} + \alpha_{10} CLOSE_{it} + \varepsilon_{it} \quad (2)$$

$$\textit{Ratio of variance of residuals of change in net income}_{it} \textit{ by variance of residuals of change in cash flow}_{it} = \left(\frac{\Delta NI^*}{\Delta CF^*}\right)_{it} \quad (3)$$

Analogous to the first metric, the *variability in ΔCF* is the variance of the residuals from a regression of the change in annual operating cash flows scaled by total assets on the control variables. The resulting second metric is the ratio of the *variability in ΔNI* and the *variability in ΔCF*. Again, a higher ratio indicates higher accounting quality.

Our third metric (metric 1.3, see *Appendix 5*) tries to capture earnings smoothing through the correlation between accruals, *ACC*, and cash flows. In line with prior studies (Ball and Shivakumar, 2005; Ball and Shivakumar, 2006; Land and Lang, 2002; Lang *et al.*, 2003; Lang *et al.*, 2006; Leuz *et al.*, 2003), we acknowledge the existence of a negative correlation between accruals and cash flows and assume that – all else equal – earnings smoothing induces a more negative correlation between the two. Leuz *et al.* (2003) further find that measures based on this correlation are highly correlated with other measures of earnings management and vary across countries according to the importance of disclosure, enforcement, equity markets, ownership concentration, and investor protection. Our third metric hence is the correlation between the residuals of the regressions from equations (4) and (5). Greater (i.e. less negative values) indicate higher accounting quality.

$$CF_{it} = \alpha_0 + \alpha_1 SIZE_{it} + \alpha_2 GROWTH_{it} + \alpha_3 EISSUE_{it} + \alpha_4 LEV_{it} + \alpha_5 DISSUE_{it} + \alpha_6 TURN_{it} + \alpha_7 AUD_{it} + \alpha_8 ADR_{it} + \alpha_9 CLOSE_{it} + \varepsilon_{it} \quad (4)$$

$$ACC_{it} = \alpha_0 + \alpha_1 SIZE_{it} + \alpha_2 GROWTH_{it} + \alpha_3 EISSUE_{it} + \alpha_4 LEV_{it} + \alpha_5 DISSUE_{it} + \alpha_6 TURN_{it} + \alpha_7 AUD_{it} + \alpha_8 ADR_{it} + \alpha_9 CLOSE_{it} + \varepsilon_{it} \quad (5)$$

where:
ACC = annual net income at fiscal year-end less annual cash flow from operating activities, scaled by end-of-year total assets (Barth *et al.*, 2008)

4.4.1.2 Management toward earnings targets

Management toward earnings targets is another form of earnings management. Prior research suggests that one common incentive to manage earnings is the avoidance of losses, and as a consequence, an unusual high frequency of small positive earnings has been detected (Burgstahler and Dichev, 1997). In order to analyze if firms manage earnings to avoid losses, we use the variable *SPOS* (*Small POSitive earnings*) which is an indicator variable equal to one if a firm's net income scaled by total assets is between 0 and 0.01 (Barth *et al.*, 2008; Lang *et al.*, 2003) and zero otherwise.[9] As dependent variable we include the indicator variable *CD* (*Cluster Differences*) which we use to test the differences in managing toward earnings targets between the clusters of our sample. Here, we adjust the approach of Barth *et al.* (2008) because we do not differentiate between IAS adopters and non-adopters but between three clusters of IFRS adopters. Likewise, we use the indicator variable *PD* (*Period Differences*) to test whether there are differences in managing towards earnings targets between the three periods for distinct clusters. Since the dependent variable is qualitative in nature, we use the logit model to regress it on the explanatory variables (Kennedy, 2008). In literature, both the probit as well as the logit model are used. The probit model uses the cumulative normal function, and the logit model uses the logistic function but there is no difference in the outcomes (Maddala, 1991). Hence, we estimate the following regression three times per time period, and thereby, compare ERC with SEC, ERC and WEC, and finally, SEC and WEC in order to see which of the respective clusters is more likely to manage toward earnings targets.

$$CD(0,1)_{it} = \alpha_0 + \alpha_1 SPOS_{it} + \alpha_2 SIZE_{it} + \alpha_3 GROWTH_{it} + \alpha_4 EISSUE_{it} + \alpha_5 LEV_{it} + \alpha_6 DISSUE_{it} + \alpha_7 TURN_{it} + \alpha_8 CF_{it} + \alpha_9 AUD_{it} + \alpha_{10} ADR_{it} + \alpha_{11} CLOSE_{it} + \varepsilon_{it} \quad (6)$$

For instance when estimating the model for the clusters SEC and WEC, *CD* takes on the value zero for observations from SEC and the value one for observations from WEC. If the coefficient of *SPOS* (which is our metric 1.4, see *Appendix 5*) would be negative, then firms from SEC

[9] The choice of this interval has not been motivated by Lang *et al.* (2003), but has been used consistently in all earnings management studies referred to in our thesis. Note that in line with Barth *et al.* (2008), we only analyze the earnings target of small positive earnings whereas Ahmed *et al.* (2012) also analyze the earnings target of the analyst consensus earnings forecast.

would be more likely to manage toward earnings targets (and hence have a lower accounting quality in this respect).

Accordingly, we also estimate the following regression three times per cluster, and thereby, compare the three periods with each other (taking the cluster as its own control) in order to see the intertemporal development of management towards earnings targets per cluster.

$$PD(0,\,1)_{it} = \alpha_0 + \alpha_1 SPOS_{it} + \alpha_2\, SIZE_{it} + \alpha_3 GROWTH_{it} + \alpha_4 EISSUE_{it} + \alpha_5 LEV_{it} + \alpha_6\, DISSUE_{it} + \alpha_7 TURN_{it} + \alpha_8 CF_{it} + \alpha_9 AUD_{it} + \alpha_{10} ADR_{it} + \alpha_{11} CLOSE_{it} + \varepsilon_{it} \quad (7)$$

For example, when estimating the model for the pre-adoption and adoption period, *PD* takes on the value zero for observations from the pre-adoption period and the value one for observations from the adoption period. If the coefficient of *SPOS* would be negative, then, firms manage toward earnings targets more frequently in the pre-adoption period than in the adoption period.

4.4.1.3 Timely loss recognition

Timely loss recognition refers to a firm's willingness to recognize large losses when they occur instead of spreading their impact over several reporting periods (Ball *et al.*, 2000). Hence, if firms manage earnings, large losses should be rare. We measure timely loss recognition as the coefficient on large negative net income, *LNNI* (*Large Negative Net Inco*me). *LNNI* is an indicator variable equal to one for a firm if its net income scaled by total assets is less than –0.2 and zero otherwise (Barth *et al.*, 2008; Lang *et al.*, 2003).[10] Similar to *4.4.1.2*, we include *CD* and *PD* as dependent variables and estimate a logit model for the three possible cluster pairs for each time period and for the three possible period combinations for each cluster, respectively.

$$CD(0,\,1)_{it} = \alpha_0 + \alpha_1 LNNI_{it} + \alpha_2 SIZE_{it} + \alpha_3 GROWTH_{it} + \alpha_4 EISSUE_{it} + \alpha_5\, LEV_{it} + \alpha_6\, DISSUE_{it} + \alpha_7 TURN_{it} + \alpha_8 CF_{it} + \alpha_9 AUD_{it} + \alpha_{10} ADR_{it} + \alpha_{11} CLOSE_{it} + \varepsilon_{it} \quad (8)$$

Analogous to *4.4.1.2*, when comparing SEC and WEC, *CD* equals zero for observations from SEC and one for WEC. A negative coefficient of *LNNI* (which is our metric 1.5, see *Appendix 5*) would indicate that firms from SEC are more likely to recognize large losses (and hence have a higher accounting quality in this respect).

$$PD(0,\,1)_{it} = \alpha_0 + \alpha_1 LNNI_{it} + \alpha_2 SIZE_{it} + \alpha_3 GROWTH_{it} + \alpha_4 EISSUE_{it} + \alpha_5\, LEV_{it} + \alpha_6\, DISSUE_{it} + \alpha_7 TURN_{it} + \alpha_8 CF_{it} + \alpha_9 AUD_{it} + \alpha_{10} ADR_{it} + \alpha_{11} CLOSE_{it} + \varepsilon_{it} \quad (9)$$

Likewise, when comparing the pre-adoption and the adoption period, *PD* equals zero for observations from the pre-adoption period and one for observations from the adoption period. A negative coefficient of *LNNI* would indicate that firms are more likely to recognize large losses in the pre-adoption period than in the adoption period.

[10] The choice of this cut off point has again not been motivated by Lang *et al.* (2003), but has been used consistently in all earnings management studies referred to in our thesis.

4.4.2 Value relevance

An accounting amount is defined as value relevant if it has a significant association with equity market values (Amir *et al.*, 1993; Barth *et al.*, 2001). Accordingly, Barth *et al.* (2008) expect that financial statements of higher accounting quality also possess a higher value relevance, expressed e.g. through a higher association of stock prices and earnings, because qualitatively higher earnings better depict a firm's underlying economic situation. For this reason, value relevance is another aspect of accounting quality. Hence, as last measures of accounting quality we include three metrics of value relevance.

In line with Barth *et al.* (2008), the first measure (metric 2.1, see *Appendix 5*) is based on the explanatory power of a regression from a firm's stock price on its net income and book value of equity. For calculating the metric, we apply a two-step approach. First, we regress stock price, *P*, on country and industry fixed effects in order to ensure that the measure of stock price is not affected by country and industry differences. Analogous to similar studies (Barth *et al.*, 2008; Lang *et al.*, 2003; Lang *et al.*, 2006), *P* is measured six months after fiscal year-end to ensure that accounting information is publicly available. As a second step, we regress the residuals from the first regression, *P**, on equity book value per share, *BVEPS*, and net income per share, *NIPS*. Finally, we take the *adjusted* R^2 of the following equation as the first value relevance metric because it captures the amount of association between stock price and earnings and equity. Higher *adjusted* R^2 indicates higher value relevance and accounting quality.

$$P_{it}^{*} = \beta_0 + \beta_1 BVEPS_{it} + \beta_2 NIPS_{it} + \varepsilon_{it} \tag{10}$$

The other two metrics of value relevance (metric 2.2 and 2.3, see *Appendix 5*) use an alternative approach to capture value relevance of financial statements. They measure the association between a firm's net income per share and its annual stock return: if firms recognize losses timely, there should be a stronger relation between contemporaneous returns and earnings due to the fact that more of the information about earnings reaches the market in the period of the loss. Both metrics are based on the same regressions, but use different sample firms to calculate the metric. Metric 2.2 uses firms with positive returns whereas metric 2.3 uses firms with negative returns. Splitting the sample into these two subsamples is based on the findings of Ball *et al.* (2000) indicating that differences in accounting quality depend on the financial performance of firms: Firms with positive returns ('good news') have lower incentives to manage their earnings. Hence, accounting quality differences are expected to be more pronounced in case of negative returns ('bad news').

Therefore, the two metrics are calculated using different sample firms, applying a two-step approach similar to the one described above: First, we regress net income per share scaled by the stock price in the beginning of the year, *NI/P*, on country and industry fixed effects. As a second step, we regress the residuals from this regression, *(NI/P)**, on annual stock return, *RETURN*. Thereby, *RETURN* is measured as the natural logarithm of the ratio of the stock price three months after fiscal year-end to the stock price nine months before fiscal year-end (adjusted for dividends and stock splits). Hence, for both subsamples, we estimate the following equation

and again use its *adjusted* R^2 as the second and third value relevance metric respectively. Again, higher *adjusted* R^2 indicates higher value relevance and accounting quality.

$$\left(\frac{NI}{P}\right)^{*}_{it} = \beta_0 + \beta_1 RETURN_{it} + \varepsilon_{it} \tag{11}$$

4.5 Accounting comparability

As discussed in *Section 2.4.3*, research struggled for a long time with the operationalization of the concept of accounting comparability. We use the empirical construct developed by De Franco *et al.* (2011) which recently gained much attention. Below, we first present this metric of accounting comparability, and then, our test design to determine the influence of the institutional setting on comparability is explained.

4.5.1 The comparability measure of De Franco *et al.* (2011)

De Franco *et al.* (2011) argue that the accounting system of any firm *i* is a "*mapping from economic events to financial statements*" (p. 899).

$$Financial\ Statements_i = f_i(Economic\ Events)_i \tag{12}$$

Hence, they argue that the accounting systems of two firms *i* and *j* are comparable if they have similar mappings, i.e. if for the same economic events they produce similar financial statements. To operationalize this concept, stock returns are taken as a proxy for the economic events a firm is exposed to in a respective period. Stock returns capture both firm-specific and industry- and even economy-wide economic events. Further, earnings are used as a proxy for financial statements output. In other words, the accounting system of firm *i* determines how stock returns (i.e. economic events) translate into reported earnings (i.e. financial statements). Thereby, first firm quarterly earnings are regressed on returns to estimate proxies of the accounting functions f_i, using four years of data for each period. One caveat to this approach is that although earnings are an important measure, it presents a limitation to use it as the only proxy of financial statements. Furthermore, stock return figures might be distorted by cross-country differences in market efficiency (Ball and Shivakumar, 2006; Holthausen, 2003). Therefore, we follow Cascino and Gassen (2012) and also regress earnings on cash flows in a second model. Accordingly, the two models are specified as follows:

$$Earnings\_MVE_{it} = \alpha_i + \beta_i Return_{it} + \varepsilon_{it} \tag{13}$$

$$Earnings\_TA_{it} = \alpha_i + \beta_i CFO_{it} + \varepsilon_{it} \tag{14}$$

Earnings_MVE are net income before extraordinary items, scaled by market value of equity at the beginning of each quarter and *Return* is the quarterly stock price return. *Earnings_TA* are net income before extraordinary items, scaled by lagged total assets and *CFO* is net cash flow

from operating activities, scaled by lagged total assets. The accounting function of firm *i* is proxied by the coefficient estimates $\hat{\alpha}_i$ and $\hat{\beta}_i$ for both models. These coefficients are truncated at the top 1 and 99 percentiles of their distributions. To estimate the similarity of accounting functions of firm *i* and *j*, next earnings for each firm in the same industry are predicted holding the economic events constant. The concept of De Franco *et al.* (2011) implies that accounting systems are comparable if for the same economic events they produce similar financial statements. Hence, for both models earnings for firms *i* and *j* are predicted for each period *t*, using the accounting functions of firms *i* and *j* but taking the same economic events (firm *i*'s return or cash flow).

$$E(Earnings\_MVE)_{iit} = \hat{\alpha}_i + \hat{\beta}_i Return_{it} \tag{15}$$

$$E(Earnings\_MVE)_{ijt} = \hat{\alpha}_j + \hat{\beta}_j Return_{it} \tag{16}$$

and

$$E(Earnings\_TA)_{iit} = \hat{\alpha}_i + \hat{\beta}_i CFO_{it} \tag{17}$$

$$E(Earnings\_TA)_{ijt} = \hat{\alpha}_j + \hat{\beta}_j CFO_{it} \tag{18}$$

Now, De Franco *et al.* (2011) define accounting comparability between firms *i* and *j* ($ACOMP_{ij}$) as

> *the negative value of the average absolute difference between the predicted earnings using firm i's and j's functions*
>
> (p. 900)

where greater (less negative) values imply greater comparability. Hence, the smaller the average earnings prediction difference, the higher is the accounting comparability between firm *i* and *j*. $CFCOMP_{ij}$ is calculated in the same way for the second model.

$$ACOMP_{ij} = -\frac{1}{16} \cdot \sum_{t-15}^{t} \left| E(Earnings\_MVE)_{iit} - E(Earnings\_MVE)_{ijt} \right| \tag{19}$$

and

$$CFCOMP_{ij} = -\frac{1}{16} \cdot \sum_{t-15}^{t} \left| E(Earnings\_TA)_{iit} - E(Earnings\_TA)_{ijt} \right| \tag{20}$$

4.5.2 Test design

A modification to the methodology of De Franco *et al.* (2011) (who analyzed a US-based setting) is required due to the international setting of our thesis. Unavailability of quarterly data and country differences in reporting frequency make it necessary to take yearly data. This is in line with the approach of Cascino and Gassen (2012) and Lang *et al.* (2010). Unlike these two studies, however, we are not interested in the consequences of IFRS adoption for adopting countries compared to non-adopting countries. Rather, our focus lies on the heterogeneity of first-order effects on different clusters of adopting countries. In other words, our aim is to investigate how the comparability of accounting information between clusters of IFRS adopting countries developed in the wake of mandatory IFRS adoption. For this reason, in contrast to Cascino and Gassen (2012) and Lang *et al.* (2010) we do not include control groups of non-adopting countries into our sample.

Hence, in our setting, we first estimate the accounting function separately for each firm for the three time periods for both models as in equations (13) and (14). Then, earnings for each firm and year are predicted with the help of the firm-period-level coefficients $\hat{\alpha}_i$ and $\hat{\beta}_i$ for both models according to equations (15), (16), (17) and (18). Then, the comparability measures are calculated as in equations (19) and (20) for each possible *i-j* combination for firms from the same industry group (based on Standard Industrial Classification with 2-digits) for each period (regardless of which cluster they are in).

In the next step, by averaging these measures $ACOMP_{i,j}$ and $CFCOMP_{i,j}$ for each cluster, peer-cluster, industry and each period, the aggregated comparability metrics for each cluster pair, industry and period are computed.

$$ACOMP_{p,ci,cj,k} = -\frac{\sum_{i,j} \left| E(Earnings\_MVE)_{iip} - E(EarningsMVE)_{ijp} \right|}{n_{p,ci,cj,k}} \tag{21}$$

$$CFCOMP_{p,ci,cj,k} = -\frac{\sum_{i,j} \left| E(Earnings\_TA)_{iip} - E(Earnings\_TA)_{ijp} \right|}{n_{p,ci,cj,k}} \tag{22}$$

$ACOMP_{p,ci,cj,k}$ and $CFCOMP_{p,ci,cj,k}$ are the aggregated accounting comparability measures over period and cluster pair and industry, *p* indicates the period, *ci* the country cluster of firm *i*, *cj* the country cluster of firm *j*, *k* the industry group and $n_{p,ci,cj,k}$ the number of firm pairs in industry *k* with firm *i* from cluster *ci* and firm *j* from cluster *cj*. At this point we have nine potential cluster pairs (including clusters with themselves) for three periods and sixty industries, ergo 1,620 results for $ACOMP_{p,ci,cj,k}$ and $CFCOMP_{p,ci,cj,k}$.

As a final step, these measures are further aggregated by averaging over all industries so that as a final metric we receive one value for the comparability between the three cluster combinations (ERC-SEC, ERC-WEC and SEC-WEC) for all three periods for both models *ACOMP* and *CFCOMP* (metric 3.1 and 3.2, see *Appendix 6*).

$$AGG.ACOMP_{p,i,j} = \frac{\sum_{i,j} ACOMP_{p,ci,cj,k}}{n_k} \tag{23}$$

and

$$AGG.CFCOMP_{p,i,j} = \frac{\sum_{i,j} CFCOMP_{p,ci,cj,k}}{n_k} \tag{24}$$

$AGG.ACOMP_{p,i,j}$ and $AGG.CFCOMP_{p,i,j}$ are the final aggregated accounting comparability measures between cluster *i* and cluster *j* in period *p*. $ACOMP_{p,ci,cj,k}$ and $CFCOMP_{p,ci,cj,k}$ are the comparability measures which are not averaged over industry yet, n_k indicates the number of industries.

As in the case of most of the accounting quality metrics, however, we are primarily not interested in the metrics *per se*, but rather whether there are significant differences between cluster pairs and periods (e.g. whether the comparability between ERC and SEC differs significantly from the comparability between ERC and WEC). Nonetheless, we cannot test for significance of these differences with a simple t-test because the empirical distribution of the metrics is unknown. Therefore, analogous to some of the accounting quality measures, we first need to generate these empirical distributions of the metrics with the help of a percentile bootstrapping approach. Thereby, we randomly select, with replacement, observations (which in this case are the 1,620 results of equations (21) and (22), $ACOMP_{p,ci,cj,k}$ and $CFCOMP_{p,ci,cj,k}$, i.e. already aggregated values for cluster pairs, industries and periods) to build representative samples that are as large as our actual sample (1,620 results per cluster pair, industry and period). Then, we calculate the metrics and the differences in each metric between clusters and time periods. By repeating this procedure 30 times we obtain the empirical distribution of these differences.[11] Based on them, we use t-tests to investigate whether the differences between clusters and periods are significant. We consider a difference significant at the $p<10\%$ ($p<5\%$, $p<1\%$) level, if a confidence interval bounded by the 5^{th} and 95^{th} (2.5^{th} and 97.5^{th}, 0.5^{th} and 99.5^{th}) percentile of the bootstrap distribution fails to contain zero.

5 Data

In this section we present the data we use to test our hypotheses. Our primary sample consists of 89,034 firm-year observations from 9,343 firms from the 24 countries for which data on enforcement characteristics was available and which adopted IFRS between 2003 and 2008 (cf. 4.1). Using Worldscope/DataStream, we obtain data beginning in 2000 and identify firms which adopted IFRS.

[11] Due to the long duration of the calculations we had to choose a lower number of replications. However, 30 replications are sufficient for meaningful results of the t-tests.

Table 3 presents a split up of firm-year observations by period and cluster for each panel. Panel 1 refers to metric 1.1 (variance in the residuals of *ΔNI* model), Panel 2 to metric 1.2 (variance of residuals from the *ΔNI* model scaled by the variance of residuals from the *ΔCF* model), Panel 3 to metric 1.3 (correlation between the residuals of the *CF* and *ACC* models), Panel 4 to metric 1.4 (differences in *SPOS* across clusters and periods), Panel 5 to metric 1.5 (differences in *LNNI* across clusters and periods), Panel 6 to metrics 2.1 to 2.3 (*adjusted R²* of regression of P*, *adjusted R²* of regression of *(NI/P)** for 'positive' subsample, *adjusted R²* of regression of *(NI/P)** for 'negative' subsample) and Panel 7 to the comparability metrics 3.1 and 3.2 (return-based and cash flow-based accounting comparability respectively). As can be seen, the number of firm-year observations differs between panels and is lower than in the primary sample. This is because we only include observations in a panel if they provide data on all variables used in the respective panel.

Panel 1 (metric 1.1)

	ERC	SEC	WEC	Total
pre-adoption	5,025	5,259	3,195	13,479
adoption	4,702	5,878	4,124	14,704
post-adoption	4,646	5,393	4,825	14,864
	14,373	16,530	12,144	43,047

Panel 2 (metric 1.2)

	ERC	SEC	WEC	Total
pre-adoption	5,004	5,239	3,100	13,343
adoption	4,704	5,878	3,524	14,106
post-adoption	4,638	5,384	4,822	14,844
	14,346	16,501	11,446	42,293

Panel 3 (metric 1.3)

	ERC	SEC	WEC	Total
pre-adoption	5,025	5,263	3,195	13,483
adoption	4,702	5,878	4,124	14,704
post-adoption	4,649	5,393	4,825	14,867
	14,376	16,534	12,144	43,054

Panel 4 (metric 1.4)

	ERC	SEC	WEC	Total
pre-adoption	5,029	5,263	3,199	13,491
adoption	4,708	5,882	4,128	14,718
post-adoption	4,649	5,393	4,831	14,873
	14,386	16,538	12,158	43,082

Panel 5 (metric 1.5)

	ERC	SEC	WEC	Total
pre-adoption	5,029	5,263	3,199	13,491
adoption	4,708	5,882	4,128	14,718
post-adoption	4,649	5,393	4,831	14,873
	14,386	16,538	12,158	43,082

Panel 6 (metrics 2.1-2.3)

	ERC	SEC	WEC	Total
pre-adoption	5,626	6,163	5,795	17,584
adoption	5,566	7,628	6,590	19,784
post-adoption	5,901	8,193	7,757	21,851
	17,093	21,984	20,142	59,219

Panel 7 (metrics 3.1 and 3.2)

	ERC	SEC	WEC	Total
pre-adoption	9,291	9,771	8,956	28,018
adoption	9,439	11,637	11,447	32,523
post-adoption	5,714	8,534	8,085	22,333
	24,444	29,942	28,488	82,874

Table 3: Number of firm-year observations by period and cluster for each panel

Table 3 shows that the clusters are more or less equally large in terms of firm-year observations although the amount of countries per cluster varies significantly, with the *SEC* cluster often comprising most firm-year observations. This is due to the fact that the ERC and *SEC* clusters – although being small in terms of the number of countries – comprise relatively large countries. Comparing the different periods, there are quite similar amounts of firm-year observations across time. Only for accounting comparability (Panel 7) there are bigger differences. Importantly, in the other panels the number of firm-year observations remains more or less

stable in the post-adoption period. However, data for 2012 was still not available in DataStream for all firms. Also, for firms whose fiscal year does not end in December (roughly 40% of our sample firms) the post-adoption period is one year shorter, i.e. comprising only three instead of four years. Normally, the post-adoption period hence would comprise fewer firm-year observations. However, in order to ensure that differences between the periods are not driven by big differences in sample sizes, for the post-adoption period we relax the restriction that data needs to be available for all four years of a period. Thus, firms for which data is only available for three years are included and consequently the post-adoption period comprises similar amounts of firm-year observations as the first two periods.

Table 4 presents a split up of firm-year observations by country for each panel. Unsurprisingly, there are quite large differences between the countries. Most observations in terms of firm-year observations come from the UK, Australia, Hong Kong, France and Singapore.

	Country	Panel 1	Panel 2	Panel 3	Panel 4	Panel 5	Panel 6	Panel 7
ERC	Finland	820	816	820	820	820	1,057	1,172
	Germany	2,780	2,769	2,783	2,791	2,791	4,687	6,227
	Netherlands	845	841	845	845	845	1,164	1,492
	Norway	651	651	651	651	651	826	1612
	United Kingdom	9,277	9,269	9,277	9,279	9,279	9,359	13,941
		14,373	14,346	14,376	14,386	14,386	17,093	24,444
	Country	**Panel 1**	**Panel 2**	**Panel 3**	**Panel 4**	**Panel 5**	**Panel 6**	**Panel 7**
SEC	Australia	5,532	5,524	5,532	5,532	5,532	9,536	13,477
	Hong Kong	5,702	5,692	5,706	5,710	5,710	5,831	7,672
	New Zealand	393	393	393	393	393	698	872
	Singapore	3,611	3,600	3,611	3,611	3,611	3,905	5,340
	South Africa	1,292	1,292	1,292	1,292	1,292	2,014	2,581
		16,530	16,501	16,534	16,538	16,538	21,984	29,942
	Country	**Panel 1**	**Panel 2**	**Panel 3**	**Panel 4**	**Panel 5**	**Panel 6**	**Panel 7**
WEC	Austria	309	305	309	309	309	497	707
	Belgium	468	448	468	468	468	758	966
	Denmark	621	629	621	629	629	895	1180
	France	3,475	3,410	3,475	3,478	3,478	4,984	6,706
	Greece	336	264	336	336	336	1,969	2,339
	Israel	505	505	505	505	505	1,612	2,852
	Italy	1,118	1,074	1,118	1,118	1,118	1,318	1,974
	Pakistan	365	365	365	365	365	779	1587
	Philippines	596	596	596	596	596	757	1114
	Portgual	208	200	208	208	208	328	447
	Spain	503	307	503	503	503	822	1053
	Sweden	1,292	1,292	1,292	1,292	1,292	2,409	3,742
	Switzerland	1,320	1,315	1,320	1,323	1,323	1,442	1,785
	Turkey	1,028	736	1,028	1,028	1,028	1,572	2,036
		12,144	11,446	12,144	12,158	12,158	20,142	28,488
	Total	43,047	42,293	43,054	43,082	43,082	59,219	82,874

Table 4: Number of firm-year observations by country for each panel

Moreover, in line with prior research we winsorize all variables of our accounting quality metrics at the 5% and 95% levels to mitigate the effects of outliers on our inferences (Barth *et*

al., 2008; Lang *et al.*, 2003).[12] *Appendix 7* presents descriptive statistics for the test and control variables used for the accounting quality metrics. As data on some variables was not available in DataStream, the number of observations per variable differs but remains on a high level. Interestingly, there are no big differences in firm size between the three clusters. However, there are differences concerning other control variables. For instance, firms from WEC have the highest leverage in all three periods. On the other hand, firms from WEC have the lowest growth rate. Growth rates in all clusters rise from the pre-adoption to the adoption period and decline again to an even lower level in the post-adoption period. Concerning the choice of auditors, there are only small differences between the clusters, but firms from *SEC* are audited slightly less often by a big audit firm. Looking at the development over time, we notice that the percentage of firms audited by a big audit company is fluctuating between approximately 55% and 63%. Untabulated tests for multicollinearity between control variables do not yield any critical results.

6 Empirical Results

In this section we present the empirical results of our thesis, first on accounting quality and then on accounting comparability, followed by a comparison with our hypotheses.

6.1 Accounting Quality

Table 5 to *Table 12* present the empirical results for our accounting quality measures. For all earnings smoothing and value relevance metrics (metrics 1.1-1.3 and 2.1-2.3 respectively) results are reported for each period-cluster pair and – more importantly, as we are interested in differences between periods and clusters – also the differences between clusters for distinct periods and the differences between periods for distinct clusters. Finally, we also report the differences-in-differences (i.e. a comparison between clusters of the magnitudes of changes in accounting quality).[13] As the managing towards earnings targets and timely loss recognition metrics (metrics 1.4 and 1.5) already compare two clusters with each other, for these metrics we only report the coefficients between clusters for distinct periods and between periods for distinct clusters. The earnings smoothing, managing towards earnings targets and timely loss recognition metrics (metrics 1.1-1.5) are jointly considered because they represent different facets of earnings management. In a second step, the results for the value relevance metrics are described.

[12] The levels were chosen in accordance with prior research. However, robustness tests (not tabulated) with variables winsorized at the 2.5% and 97.5% level do not change the inferences from our results.

[13] To illustrate the differences-in-differences take a look at *Table 5*: It can be seen that accounting quality as defined by the variance of net income metric declines for firms from both the ERC and the SEC cluster from the first to the second period (from 0.0193 to 0.0127 for firms from ERC and from 0.0196 to 0.0178 for firm from SEC). Hence, the difference between the two periods is 0.0066 for ERC and 0.0018 for SEC. The difference-in-difference here compares the differences between the two periods of ERC and SEC, thus comparing the magnitude of changes in accounting quality relative to the other cluster. In this case it is 0.0047 (~0.0066-0.0018), indicating that accounting quality decreased stronger for firms from ERC than for firms from SEC and that the gap in accounting quality between the two clusters increased.

6.1.1 Description of results

6.1.1.1 Earnings management

In order to find patterns between the results of the five metrics testing for earning management (for a summary of the metrics see *Appendix 5*), we compare the results using three dimensions: First, the clusters are ranked in terms of earnings management (i.e. ranking the clusters by the level of earnings management for each period). Second, the development over time for each cluster is analyzed (i.e. the change in the level of earnings management for each cluster from period to period). Finally, the differences-in-differences are used to evaluate whether the level of earnings management changes with different magnitudes in different clusters. We do this for each of the three time periods comparing also the changes between the periods. From these analyses of earnings management practices, we draw conclusions for accounting quality in general.

Change from pre-adoption to adoption period

Starting with the pre-adoption period, the results indicate that firms from WEC show significantly lower values for all five earnings management measures compared to firms from ERC and SEC (lower values indicating lower accounting quality or higher levels of earnings management for all five metrics). Taking the variability of change in net income (metric 1.1, *Table 5*) as an example, the variability is 0.0097 for WEC, 0.0193 for ERC and 0.0196 for SEC. Hence, while the difference between ERC and WEC is 0.0095 and 0.0098 for SEC and WEC respectively, indicating that firms from WEC have the lowest accounting quality in the first period as expressed by this metric.

Taking timely loss recognition (metric 1.5, see *Table 9*) as an example of a relative metric, the coefficient of *LNNI* between ERC and WEC is -0.4198 and -0.4583 between SEC and WEC. Negative values indicate here that firms from the first-mentioned cluster (e.g. in case of ERC vs. WEC it is ERC) have a higher likelihood to recognize large negative losses and hence a higher accounting quality as they recognize occurred losses immediately and do not try to spread them over longer periods. Summarizing these results, firms from WEC exhibit the highest level of earnings management in the pre-adoption period.

Comparing ERC and SEC, the cash flow-accrual correlation metric (metric 1.3) and the *LNNI* metric (metric 1.5) imply that firms from SEC have a significantly higher accounting quality than firms from ERC. For the other three metrics, no significant differences between these two clusters can be found. Hence, we assume that both clusters show fairly similar levels of earnings management in the pre-adoption period. Now that the ranking of the clusters for the pre-adoption period regarding earnings management and, correspondingly, accounting quality is known, we proceed ranking the clusters for the second period.

			Cluster(s)						
			Absolute values			Differences			
			ERC	SEC	WEC	ERC-SEC	ERC-WEC	SEC-WEC	
Period(s)	Absolute values	1	0.0193	0.0196	0.0097	-0.0003	0.0095***	0.0098***	
		2	0.0127	0.0178	0.0064	-0.0051***	0.0064***	0.0114***	
		3	0.0147	0.0176	0.0078	-0.003***	0.0069***	0.0099***	
	Differences	(1)-(2)	0.0066***	0.0018**	0.0034***	0.0047***	0.0032***	-0.0016***	Difference-in-differences
		(1)-(3)	0.0046***	0.002***	0.002***	0.0026***	0.0026***	0.0001***	
		(2)-(3)	-0.002**	0.0002*	-0.0014***	-0.0021***	-0.0006***	0.0016***	
							Differences-in-differences		

This table presents the results of the first earnings management metric - the variance of the residuals from a regression of the change in net income on the control variables from equation 1 - for each period and cluster (upper-left segment), the differences between clusters for distinct periods (upper-right segment), the differences between periods for distinct clusters (lower-left segment) and the differences-in-differences between periods and clusters (lower-right segment). The differences-in-differences compare the differences between periods over clusters. We number the periods for ease of presentation where period 1 is the pre-adoption period, period 2 the adoption period and finally, period 3 the post-adoption period. Values are rounded to four decimal places. We test significance using a percentile-bootstrapping procedure. Thus, we get the empirical distribution of the differences of the metric. *, ** and *** indicate significance at the p<10%, p<5% and p<1% level, respectively.

Table 5: Metric 1.1 Variance of *ΔNI (Panel 1)**

			Cluster(s)						
			Absolute values			Differences			
			ERC	SEC	WEC	ERC-SEC	ERC-WEC	SEC-WEC	
Period(s)	Absolute values	1	1.9504	1.9946	1.5768	-0.0442	0.3736***	0.4178***	
		2	1.6177	1.9450	1.4845	-0.3273***	0.1332*	0.4604***	
		3	1.7851	1.9486	1.2910	-0.1636	0.4941***	0.6577***	
	Differences	(1)-(2)	0.3327***	0.0497	0.0923	0.2831***	0.2404***	-0.0426***	Difference-in-differences
		(1)-(3)	0.1653***	0.046**	0.2858***	0.119***	-0.1205*	-0.2399***	
		(2)-(3)	-0.1674	-0.0037	0.1936	-0.1637***	-0.3609***	-0.1972	
							Differences-in-differences		

This table presents the results of the second earnings management metric - the ratio from equation 3 between the variance of the residuals from a regression of the change in net income on the control variables and the variance of the residuals from a regression of the change in cash flow on the control variables - for each period and cluster (upper-left segment), the differences between clusters for distinct periods (upper-right segment), the differences between periods for distinct clusters (lower-left segment) and the differences-in-differences between periods and clusters (lower-right segment). The differences-in-differences compare the differences between periods over clusters. We number the periods for ease of presentation where period 1 is the pre-adoption period, period 2 the adoption period and finally, period 3 the post-adoption period. Values are rounded to four decimal places. We test significance using a percentile-bootstrapping procedure. Thus, we get the empirical distribution of the differences of the metric. *, ** and *** indicate significance at the p<10%, p<5% and p<1% level, respectively.

Table 6: Metric 1.2 Variance of *ΔNI*/ΔCF (Panel 2)**

			Cluster(s)						
			Absolute values			Differences			
			ERC	SEC	WEC	ERC-SEC	ERC-WEC	SEC-WEC	
Period(s)	Absolute values	1	-0.1313	-0.0556	-0.2827	-0.0757**	0.1513***	0.2271***	
		2	-0.1551	-0.1449	-0.3960	-0.0102	0.2409***	0.2511***	
		3	-0.1438	-0.2403	-0.4138	0.0965***	0.27***	0.1735***	
	Differences	(1)-(2)	0.0237	0.0893***	0.1133***	-0.0656***	-0.0896***	-0.0241***	Difference-in-differences
		(1)-(3)	0.0124	0.1847***	0.1311***	-0.1723***	-0.1187***	0.0536***	
		(2)-(3)	-0.0113	0.0954***	0.0178	-0.1067***	-0.0291	0.0776***	
						Differences-in-differences			

This table presents the results of the third earnings management metric - the correlation between the residuals from a regression of cash flows on the control variables from equation 4 and the residuals from a regression of accruals on the control variables from equation 5 - for each period and cluster (upper-left segment), the differences between clusters for distinct periods (upper-right segment), the differences between periods for distinct clusters (lower-left segment) and the differences-in-differences between periods and clusters (lower-right segment). The differences-in-differences compare the difference between periods over clusters. We number the periods for ease of presentation where period 1 is the pre-adoption period, period 2 the adoption period and finally, period 3 the post-adoption period. Values are rounded to four decimal places. We test significance using a percentile-bootstrapping procedure. Thus, we get the empirical distribution of the differences of the metric. *, ** and *** indicate significance at the p<10%, p<5% and p<1% level, respectively.

Table 7: Metric 1.3 Correlation between accruals and cash flows (Panel 3)

			Cluster(s)					
						Coefficients		
			ERC	SEC	WEC	ERC-SEC	ERC-WEC	SEC-WEC
Period(s)		1				-0.0480	0.3685***	0.4253***
		2				0.0311	0.2984***	0.3062***
		3				-0.2015*	0.1748	0.4607***
	Coefficients	(1)-(2)	-0.2177**	-0.1499	-0.2925***			
		(1)-(3)	0.0482	-0.1649*	-0.0630			
		(2)-(3)	0.2848**	0.0011	0.1991*			

This table presents the results of the fourth earnings management metric - the coefficient on small positive earnings (SPOS) in the regressions given in equations 6 and 7. In the upper-right segment there are the results of the metric between clusters for distinct periods, in the lower-left segment the results between periods for distinct clusters. Negative values imply that in the cluster comparisons (e.g. ERC-SEC) firms from the first-mentioned cluster (i.e. here ERC firms) are more likely to report small positive earnings and that in the period comparisons (e.g. 1-2) firms are more likely to report small positive earnings in the first-mentioned period (i.e. here in the first period). We number the periods for ease of presentation where period 1 is the pre-adoption period, period 2 the adoption period and finally, period 3 the post-adoption period. Values are rounded to four decimal places. We test significance using a Wald-test, the values for the Wald-statistics are not tabulated. *, ** and *** indicate significance at the p<10%, p<5% and p<1% level, respectively.

Table 8: Metric 1.4 Small positive earnings (Panel 4)

			Cluster(s)					
						Coefficients		
			ERC	SEC	WEC	ERC-SEC	ERC-WEC	SEC-WEC
Period(s)		1				0.1754*	-0.4198***	-0.4583***
		2				0.5029***	-0.5188***	-1.087***
		3				0.3254**	-0.4053***	-0.6271***
	Coefficients	(1)-(2)	-0.5514***	-0.1607*	-0.7577***			
		(1)-(3)	-0.5099***	-0.3552***	-0.5819***			
		(2)-(3)	0.0585	-0.1792***	0.1187			

This table presents the results of the fifth earnings management metric - the coefficient on large negative net income (LNNI) in the regressions given in equations 8 and 9. In the upper-right segment there are the results of the metric between clusters for distinct periods, in the lower-left segment the results between periods for distinct clusters. Negative values imply that in the cluster comparisons (e.g. ERC-WEC) firms from the first-mentioned cluster (i.e. here ERC firms) are more likely to recognize large losses and that in the period comparisons (e.g. 1-2) firms are more likely to recognize large losses in the first-mentioned period (i.e. here in the first period). We number the periods for ease of presentation where period 1 is the pre-adoption period, period 2 the adoption period and finally, period 3 the post-adoption period. Values are rounded to four decimal places. We test significance using a Wald-test, the values for the Wald-statistics are not tabulated. *, ** and *** indicate significance at the p<10%, p<5% and p<1% level, respectively.

Table 9: Metric 1.5 Large negative net income (Panel 5)

Here again, all five metrics significantly imply that WEC shows the highest levels of earnings management. For example, the coefficients for small positive earnings (metric 1.4, see *Table* 8) are 0.2984 between ERC and WEC and 0.3062 between SEC and WEC. Positive values imply that firms from the last-named cluster have a higher tendency to report small positive earnings (i.e. in this case for both comparisons firms from WEC) and hence manage more often earnings which in turn leads to lower accounting quality. Comparing ERC and SEC, again there are mixed results with two metrics (namely, small positive earnings, metric 1.4, and correlation between cash flow and accruals, metric 1.3) indicating no significant differences and the other three metrics indicating that firms from SEC show lower levels of earnings management in the adoption period than firms from ERC. Therefore, again it is argued that SEC and ERC firms exhibit fairly similar levels of earnings management. Knowing the relative ranking of the three clusters in the first two periods, as a next step it is analyzed whether earnings management increased or decreased between the first and second period in the three clusters.

The findings show that for most of the five metrics the values for all clusters are highest in the pre-adoption period. Taking the variance of change in net income (metric 1.1, see *Table 5*) as an example, we find that ERC has a value of 0.0193 in the first and 0.0127 in the second period (meaning an increase in earnings management and a decrease of accounting quality by 0.0066). For SEC, it is a decrease in accounting quality of 0.0018 and for WEC a decrease of 0.0034 respectively. For metrics 1.2 and 1.3 as well as for the coefficients of large negative net income (metric 1.5) the tenor is the same, i.e. that accounting quality is highest in the pre-adoption period (only a minority of values is not significant). Only the coefficients of small positive earnings (metric 1.4, see *Table 8*) indicate that accounting quality is higher in the second or third period than in the pre-adoption period. For example, the coefficient for ERC between the

first and the second period is -0.2177. Negative values again imply that firms are more likely to manage towards earnings targets, and consequently, that accounting quality is lower in the first-mentioned period. In other words, this metric implies that accounting quality is higher in the adoption period than in the pre-adoption period. However, as all other metrics indicate the opposite we conclude that accounting quality generally decreases for all clusters from the pre-adoption to the adoption period.

Hence, so far the results allow a relative ranking of the clusters in the first two periods and indicate a decrease in accounting quality in the second period for all clusters compared to the first period. Therefrom arises the question whether the decrease in accounting quality was equally strong for all clusters or whether some clusters experienced stronger decreases than others. The differences-in-differences analysis can answer this question as it measures differences in the magnitude of accounting quality changes between the three clusters. If the differences-in-differences are statistically not significant, it would imply that the differences between the clusters remain stable over time and that all clusters follow the same overall trend (i.e. experience a change in accounting quality of similar magnitude). However, if the differences-in-differences are significantly different from zero, this would mean that the clusters experience changes in accounting quality of different magnitude. This again would imply that the clusters follow different underlying trends and the differences in accounting quality between the clusters are changing over time, meaning that the clusters are either converging or drifting apart.

As we did not calculate differences-in-differences for the managing towards earnings targets and timely loss recognition metrics (metrics 1.4 and 1.5) due to the different construction of the metrics, only the three earnings smoothing metrics (metrics 1.1 to 1.3) are considered in the analysis of the differences-in-differences. Here, we only include results into our interpretation if all three differences-in-differences are consistent and show the same direction. For example, if two metrics show that accounting quality decreases stronger for firms from WEC than for firms from ERC between the first and the second period, but if the third metric shows that accounting quality decreases stronger for firms from ERC than for firms from WEC, then we would not use these differences-in-differences for our inferences. Put differently, in that case no conclusions about different magnitudes of accounting quality changes would be drawn due to the opposed findings of the metrics, but we would only acknowledge the existence of other factors which affect the clusters in different ways and cannot be omitted.

Taking a look at the figures, it becomes clear that only for the differences-in-differences between SEC and WEC from the pre-adoption to the adoption period the results show a consistent picture. Looking at the variance of change in net income over variance of change in cash flow metric (metric 1.2, see Table 6) as an example, it can be seen that for firms from SEC the metric decreased from 1.9946 to 1.9450 from the pre-adoption to the adoption period (i.e. a decline of 0.0497). For WEC the decrease equals 0.0923. Thus, the difference-in-difference between SEC and WEC from the first to the second period is -0.0426, meaning that accounting quality decreased stronger for WEC firms than for SEC firms. The differences-in-differences for the other two earnings smoothing measures (metrics 1.1 and 1.3, see *Table 5* and *Table 7*)

show the same results (stronger decrease for WEC than for SEC). However, for the other two cluster pairs (SEC-ERC, ERC-WEC) there are no consistent differences-in-differences.

Summing up so far, accounting quality in terms of earnings management decreased for all three clusters from the pre-adoption to the adoption period and in both periods accounting quality was lowest for firms from WEC. SEC and ERC exhibit rather similar accounting quality in both periods, SEC probably being slightly higher. Further, we know that accounting quality decreased stronger for firms from WEC than for firms from SEC between the first two periods which is consistent with more opportunistic use of increased accounting discretion by WEC firms. Next, the developments in the post-adoption period are described. The focus lies on whether accounting quality increased or decreased compared to the first two periods.

Change from adoption to post-adoption period

First, the relative ranking between the clusters in the post-adoption period is analyzed. Similar results for WEC are obtained for four metrics indicating that firms from WEC show the highest level of earnings management (only the coefficient of small positive earnings, metric 1.4, shows no significant differences between ERC and WEC). Therefore, WEC shows also in the post-adoption period the lowest accounting quality of all clusters. Comparing ERC and SEC, there are three metrics indicating that accounting quality of firms from SEC is higher than for firms from ERC. The variance of change in net income over variance of change in cash flow (metric 1.2, see *Table 6*) indicates that there is no significant difference between the two clusters. Finally, the correlation between cash flow and accruals (metric 1.3, see *Table 7*) implies that ERC has a higher accounting quality. For this reason, it can be stated that accounting quality is approximately at the same level for ERC and SEC firms in the post-adoption period. The next question is whether accounting quality increased or decreased further from the adoption to the post-adoption period.

Taking the first earnings smoothing metric (metric 1.1, see *Table 5*) as an example, we find that ERC has a value of 0.0127 in the second and 0.0147 in the third period (indicating a decrease in earnings management and hence an increase in accounting quality by 0.002). For SEC, it is an increase in accounting quality of 0.0002 (though at a smaller level of significance) and for WEC also a decrease of 0.0014 from the second to the third period. Unfortunately, not all metrics yield similar results. For firms from ERC there are three metrics indicating that accounting quality did not change significantly between the second and the third period. The variance of change in net income (metric 1.1, see *Table 5*), however, indicates that accounting quality is higher in the third period than in the second period (i.e. increased again). Lastly, the coefficient on small positive earnings (metric 1.4, see *Table 8*) implies that accounting quality is higher in the second than the third period (i.e. a further decrease). For SEC firms, three metrics imply that accounting quality decreased further in the post-adoption period and two metrics (metrics 1.2 and 1.4, see *Table 6* and *Table 8* respectively) show no significant differences between the second and the third period. For WEC firms, three metrics show no significant differences, one (metric 1.1, see *Table 5*) implies that accounting quality increased again and another (metric 1.4, see *Table 8*) that accounting quality decreased further. Due to these inconsistent results across different metrics it is not possible to identify the exact change of accounting quality between the adoption and post-adoption period. Instead, only ranges of

potential outcomes or developments can be derived. For firms from ERC and WEC, accounting quality could have remained stable, increased or decreased. For firms from SEC, accounting quality remained stable or decreased.

Hence, we cannot pinpoint the exact development of accounting quality in the post-adoption period. However, we find that accounting quality generally is highest in the pre-adoption period (except for in the managing towards earnings targets metric 1.4). Comparing the accounting quality levels of the pre- and post-adoption period, we can therefore conclude that even in the case that accounting quality increased in the post-adoption period in some clusters it will still be at a significantly lower level than in the pre-adoption period.

Figure 3 summarizes our empirical results of the development of accounting quality (as indicated by the levels of earnings management) over clusters and periods. It shows that accounting quality differs between firms from WEC on the one hand and firms from ERC or SEC on the other hand in all three periods. Between firms from SEC and ERC generally no significant differences in accounting quality can be noticed, but firms from SEC rather showing a slightly higher accounting quality. Regarding the development over time, accounting quality decreased for all clusters from the pre-adoption to the adoption period with the decrease being strongest for firms from WEC. Concerning the post-adoption period, no unambiguous inference can be made. Rather, accounting quality might have further decreased, remained stable or (except for firms from SEC) increased again. However, accounting quality definitely is at a lower level than in the pre-adoption period for all three clusters, even in case of an increase in the last period. Due to the ambiguous findings in the last period, in *Figure 3* the shaded areas depict a range of potential outcomes of accounting quality. For ease of presentation there is only one shaded area for both ERC and SEC (otherwise there would be two partially overlapping areas). Finally, the results can be split up into different facets of earnings management. While the summary results described so far refer to the level of earnings smoothing and timely loss recognition, the managing towards earnings targets metric (metric 1.4) rather indicates a (temporary) improvement in accounting quality. Nevertheless, as this is the only metric showing differing results, we will focus in the later analysis on the implications of the majority of the earnings management metrics.

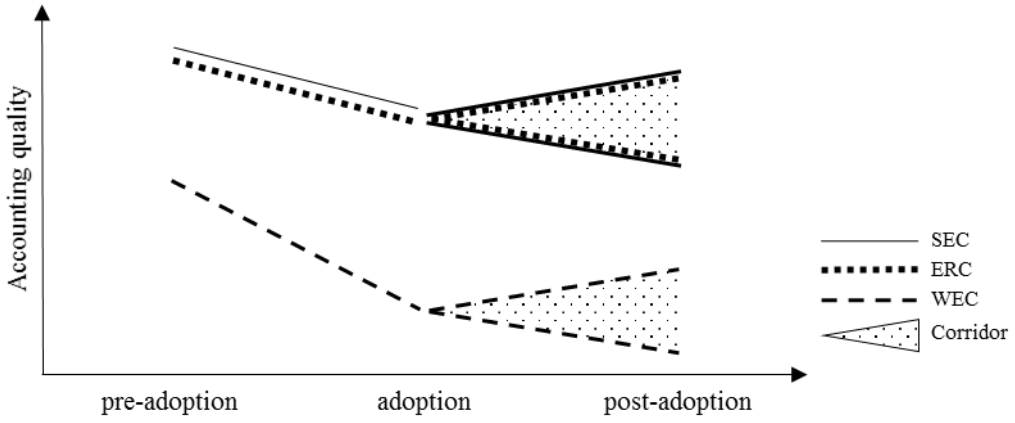

Figure 3: Accounting quality in terms of earnings management

6.1.1.2 Value Relevance

In this section we analyze the results of the tests of the value relevance metrics (for a summary of the metrics see *Appendix 5*). However, the results of these metrics draw an inconsistent picture and show less often statistical significance. For example, the results of the second value relevance metric (metric 2.2, see *Table 11*) do not even show one significant difference, neither between clusters for distinct periods nor between periods for distinct clusters. Furthermore analyzing the relative ranking of the country clusters for each period, the first value relevance metric (metric 2.1, see *Table 10*) often shows diametrically opposite results compared to the third metric (metric 2.3, see *Table 12*). For instance, for the pre-adoption period metric 2.1 indicates that firms from WEC have higher value relevance than firms from SEC and that there are no significant differences compared to ERC firms (see *Table 10*). Metric 2.3, on the other hand, indicates that for the first period firms from SEC show higher value relevance compared to both firms from ERC and WEC (see *Table 12*). At the same time, firms from the two latter clusters do not show significant differences in value relevance.

			Cluster(s)						
			Absolute values			Differences			
			ERC	SEC	WEC	ERC-SEC	ERC-WEC	SEC-WEC	
Period(s)	Absolute values	1	4.69%	-0.01%	7.39%	4.70%	-2.70%	-7.4%**	
		2	15.06%	1.41%	19.79%	13.65%**	-4.73%	-18.38%***	
		3	0.48%	6.03%	36.42%	-5.54%	-35.94%***	-30.39%***	
	Differences	(1)-(2)	-10.37%	-1.42%	-12.4%**	-8.95%***	2.03%**	10.98%***	Difference-in-differences
		(1)-(3)	4.21%	-6.04%	-29.03%***	10.25%***	33.24%***	22.99%***	
		(2)-(3)	14.58%	-4.62%	-16.63%**	19.2%***	31.21%***	12.02%***	
							Differences-in-differences		

This table presents the results of the first value relevance metric - the adjusted R^2 from a regression of stock price (after country- and industry-fixed-effects) on the equity book value per share and net income per share - for each period and cluster (upper-left segment), the differences between clusters for distinct periods (upper-right segment), the differences between periods for distinct clusters (lower-left segment) and the differences-in-differences between periods and clusters (lower-right segment). The differences-in-differences compare the differences between periods over clusters. Note that a negative value for adjusted R^2 can be possible because unlike the normal R^2 adjusted R^2 does not have the illustrative meaning of how much of the variation can be explained by the model. The numbers in the upper-left segment are percentage numbers, the others differences (or differences-in-differences) in percentage points. Values are rounded to two decimal places. We number the periods for ease of presentation where period 1 is the pre-adoption period, period 2 the adoption period and finally, period 3 the post-adoption period. We test significance using a percentile-bootstrapping procedure. Thus, we get the empirical distribution of the differences of the metric. *, ** and *** indicate significance at the p<10%, p<5% and p<1% level, respectively.

Table 10: Metric 2.1 Regression of P* - *adjusted R^2* (Panel 6)

			Cluster(s)						
			Absolute values			Differences			
			ERC	SEC	WEC	ERC-SEC	ERC-WEC	SEC-WEC	
Period(s)	Absolute values	1	1.84%	1.16%	1.73%	0.68%	0.11%	-0.5680%	
		2	3.08%	1.84%	0.57%	1.2336%	2.51%	1.2728%	
		3	2.85%	1.72%	4.02%	1.12%	-1.1673%	-2.2918%	
	Differences	(1)-(2)	-1.24%	-0.68%	1.1583%	-0.5527%	-2.39%***	-1.84%***	Difference-in-differences
		(1)-(3)	-1.01%	-0.56%	-2.2867%	-0.4436%	1.2802%	1.7239%	
		(2)-(3)	0.23%	0.12%	-3.4450%	0.1091%	3.67%***	3.56%***	
									Differences-in-differences

This table presents the results of the second value relevance metric - the adjusted R^2 from a regression of net income per share (after country- and industry-fixed-effects) on annual stock return for firms with a positive stock return (good news) - for each period and cluster (upper-left segment), the differences between clusters for distinct periods (upper-right segment), the differences between periods for distinct clusters (lower-left segment) and the differences-in-differences between periods and clusters (lower-right segment). The differences-in-differences compare the differences between periods over clusters. Note that a negative value for adjusted R^2 can be possible because unlike the normal R^2 adjusted R^2 does not have the illustrative meaning of how much of the variation can be explained by the model. The numbers in the upper-left segment are percentage numbers, the others differences (or differences-in-differences) in percentage points. Values are rounded to two decimal places. We number the periods for ease of presentation where period 1 is the pre-adoption period, period 2 the adoption period and finally, period 3 the post-adoption period. We test significance using a percentile-bootstrapping procedure. Thus, we get the empirical distribution of the differences of the metric. *, ** and *** indicate significance at the p<10%, p<5% and p<1% level, respectively.

Table 11: Metric 2.2 Regression of (NI/P)* - *adjusted R^2* good news (Panel 6)

			Cluster(s)						
			Absolute values			Differences			
			ERC	SEC	WEC	ERC-SEC	ERC-WEC	SEC-WEC	
Period(s)	Absolute values	1	0.20%	1.36%	-0.03%	-1.16%*	0.23%	1.39%*	
		2	1.79%	0.03%	0.92%	1.76%**	0.87%	-0.8832%	
		3	1.09%	-0.01%	0.01%	1.11%	1.0841%	-0.0213%	
	Differences	(1)-(2)	-1.59%*	1.33%**	-0.9443%	-2.92%***	-0.65%***	2.27%***	Difference-in-differences
		(1)-(3)	-0.89%	1.37%**	-0.0341%	-2.27%***	-0.86%***	1.41%***	
		(2)-(3)	0.70%	0.05%	0.9102%	0.65%*	-0.21%**	-0.86%***	
									Differences-in-differences

This table presents the results of the third value relevance metric - the adjusted R^2 from a regression of net income per share (after country- and industry-fixed-effects) on annual stock return for firms with a negative stock return (bad news) - for each period and cluster (upper-left segment), the differences between clusters for distinct periods (upper-right segment), the differences between periods for distinct clusters (lower-left segment) and the differences-in-differences between periods and clusters (lower-right segment). The differences-in-differences compare the differences between periods over clusters. Note that a negative value for adjusted R^2 can be possible because unlike the normal R^2 adjusted R^2 does not have the illustrative meaning of much of the variation can be explained by the model. The numbers in the upper-left segment are percentage numbers, the others differences (or differences-in-differences) in percentage points. Values are rounded to two decimal places. We number the periods for ease of presentation where period 1 is the pre-adoption period, period 2 the adoption period and finally, period 3 the post-adoption period. We test significance using a percentile-bootstrapping procedure. Thus, we get the empirical distribution of the differences of the metric. *, ** and *** indicate significance at the p<10%, p<5% and p<1% level, respectively.

Table 12: Metric 2.3 Regression of (NI/P)* - *adjusted R^2* bad news (Panel 6)

Similarly, metric 2.1 and 2.3 show insignificant changes over time periods per cluster in three out of six cases. Hence, it cannot be inferred how value relevance developed over time for each of the country clusters. With both the relative ranking and the development over time being ambiguous, an analysis of differences-in-differences is difficult to interpret. It can only be noted that again the occurrence of consistent differences-in-differences over all three value relevance metrics indicates the existence of other factors which have different effects for different clusters without being able to concretize this.

To sum up, it is not possible to draw any meaningful conclusions about the development of value relevance over clusters and periods. For this reason, it is neither possible to draw a graph concerning the development of the value relevance of financial statements analogous to *Figure 3*.

6.1.2 Evaluation of hypotheses

Based on the empirical results, we next evaluate our hypotheses concerning accounting quality. As no inferences can be made about value relevance, only the results on earnings management are used to draw conclusions about accounting quality. Therefore, in the following evaluation of our hypotheses and in the discussion of the results accounting quality always refers to accounting quality in terms of earnings management.

Based on the findings of prior literature we expected in H1 that firms from countries which orchestrated concurrent enforcement reforms together with the adoption of IFRS (ERC) show an increase in accounting quality because the improvement in their enforcement setting leads to an increase of reporting incentives for firms. Second, we expected in H2 that firms from countries that already had relatively strong enforcement systems prior to IFRS adoption (most of the countries having an Anglo-Saxon background and thus having a greater proximity to IFRS than some continental European countries) to show no significant increase in accounting quality because reporting incentives were already strong before IFRS adoption. Finally, we hypothesized in H3 that firms from countries with relatively weak enforcement (WEC) will experience no increase in accounting quality through the adoption of IFRS because they have no strong incentives to report high quality financial statements.

Looking at *Figure 1*, it can be stated that H1 needs to be rejected as accounting quality of firms from ERC did not increase but decreased. In contrast, H2 and H3 are confirmed in that SEC and WEC firms indeed did not show an increase in quality. However, based on our expectation of a more or less steady development of accounting quality of these two latter clusters, the actual decreases are to some extent surprising. Therefore, in our discussion in *Section* 7 the implications of our results, e.g. concerning the role of enforcement, are discussed in more detail.

6.2 Accounting comparability

6.2.1 Description of results

Table 13 and *Table 14* present the empirical results for the aggregated accounting comparability metrics (metrics 3.1 and 3.2, see *Appendix 6* for a summary of their calculation) based on De Franco *et al.* (2011). They show one aggregated value of accounting comparability for each

cluster combination for each period in the upper left segment (with higher, i.e. less negative, values indicating higher comparability). In the upper right segment, the differences between two cluster pairs for each period are reported (e.g. whether accounting comparability between firms from ERC and SEC is higher than between firms from ERC and WEC). In the lower left segment, the differences between periods for each cluster pair are calculated (e.g. whether accounting comparability between firms from ERC and SEC is higher in the pre-adoption than in the adoption period). The differences-in-differences are reported in the lower right segment of *Table 13* and *Table 14*. Analogous to the earnings smoothing and value relevance metrics (metrics 1.1-1.3 and 2.1-2.3), the differences-in-differences indicate whether accounting comparability between two cluster pairs changes with different magnitudes between two periods (e.g. whether accounting comparability between firms from ERC and SEC decreased stronger than between firms from ERC and WEC from the first to the second period). Besides the tables, *Figure 4* and *Figure 5* depict the development of accounting comparability for each cluster and period for the two metrics.

6.2.1.1 *AGG.ACOMP* metric 3.1

Starting with the return-based *AGG.ACOMP* measure of accounting comparability (metric 3.1, see *Table 13* and *Figure 4*), the results indicate that in the pre-adoption period accounting comparability between firms from ERC and SEC (-0.3950) is significantly higher than between firms from ERC and WEC (-0.4358) as well as between firms from SEC and WEC (-0.4406). The difference in accounting comparability between firms from ERC and WEC and between firms from SEC and WEC is not significant. Having established the relative ranking of the cluster pairs in terms of accounting comparability for the first period, the same is done for the second period as a next step. In the adoption period accounting comparability between firms from ERC and WEC (-0.2677) is significantly higher than between firms from ERC and SEC (-0.3148) as well as between firms from SEC and WEC (-0.3237). In this period, the difference in accounting comparability between firms from ERC and SEC and firms from SEC and WEC is not significant. Hence, the relative ranking of the cluster pairs changed from the pre-adoption to the adoption period. Consequently, accounting comparability between the cluster pairs must have changed with different magnitudes. Before analyzing the differences-in-differences, though, the change in accounting comparability from the first to the second period is examined for each cluster.

The findings show that accounting comparability according to metric 3.1 increased for all the three cluster pairs:[14] between firms from ERC and SEC by -0.0802 (-0.3950-(-0.3148)), between firms from ERC and WEC by -0.1681 as well as between firms from SEC and WEC by -0.1168. Furthermore, the increase between firms from ERC and WEC has been significantly stronger than between firms from the other two cluster pairs. In *Figure 4*, this corresponds to the stronger incline of the line representing the comparability between firms from ERC and WEC. Regarding the other two cluster pairs, accounting comparability significantly increased stronger between firms from SEC and WEC than between firms from ERC and SEC.

[14] Note that an increase in accounting comparability corresponds to a negative difference between periods as less negative values imply greater accounting comparability.

			Cluster pair(s)						
			Differences						
			ERC vs. SEC	ERC vs. WEC	SEC vs. WEC	(ERC vs. SEC) - (ERC vs. WEC)	(ERC vs. SEC) - (SEC vs. WEC)	(ERC vs. WEC) - (SEC vs. WEC)	
Period(s)	Differences	1	-0.3950	-0.4358	-0.4406	0.0408***	0.0456***	0.0048***	
		2	-0.3148	-0.2677	-0.3237	-0.0471***	0.0089***	0.0561***	
		3	-0.4971	-0.4718	-0.4796	-0.0253***	-0.0175***	0.0078***	
		(1)-(2)	-0.0802***	-0.1681***	-0.1168***	0.0879***	0.0366***	-0.0513***	Difference-in-differences
		(1)-(3)	0.1021***	0.0360***	0.0390***	0.0661***	0.0631***	-0.003*	
		(2)-(3)	0.1823***	0.2041***	0.1558***	-0.0218***	0.0264***	0.0482***	
						Differences-in-differences			

The table presents the results of the first, return-based accounting comparability metric from equation 23 for the aggregated average accounting comparability between two clusters for each period (upper-left segment), the differences between two cluster pairs for distinct periods (upper-right segment), the differences between periods for distinct cluster pairs (lower-left segment) and the differences-in-differences between periods and cluster pairs (lower-right segment). The differences-in-differences compare the differences between periods over cluster pairs. We number the periods for ease of presentation where period 1 is the pre-adoption period, period 2 the adoption period and finally, period 3 the post-adoption period. Values are rounded to four decimal places. We test significance using a percentile-bootstrapping procedure. Thus, we get the empirical distribution of the differences of the metric. *, ** and *** indicate significance at the p<10%, p<5% and p<1% level, respectively.

Table 13: Metric 3.1 *AGG.ACOMP* (Panel 7)

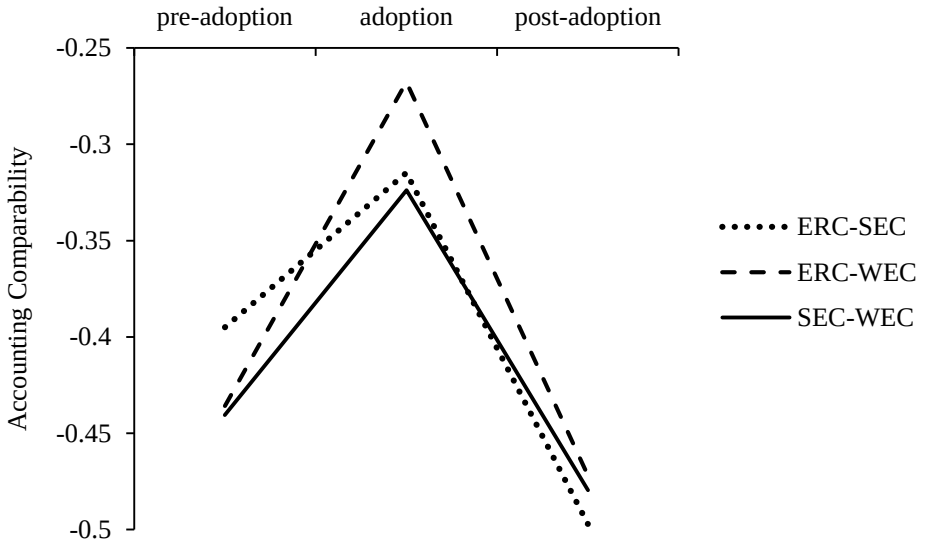

Figure 4: Return-based accounting comparability metric

In the post-adoption period the differences in accounting comparability between the three cluster pairs reduced: accounting comparability between firms from ERC and WEC (-0.4718) does not significantly differ from the accounting comparability between firms from SEC and WEC (-0.4796). Neither does the accounting comparability between firms from SEC and WEC significantly differ from the accounting comparability between firms from ERC and SEC (-0.4791). Only the accounting comparability between firms from ERC and WEC is significantly higher than between firms from ERC and SEC, but with a small level of

significance. Hence, accounting comparability between all three cluster pairs converged in the post-adoption period.

Regarding the change in accounting comparability between the adoption and the post-adoption period, comparability decreases for all three cluster pairs: between firms from ERC and SEC by 0.1823, between firms from ERC and WEC by 0.2041 as well as between firms from SEC and WEC by 0.1558. The decrease is significantly stronger for the comparability between firms from ERC and WEC than for the comparability between firms from SEC and WEC. Accounting comparability also significantly decreases stronger between firms from ERC and WEC than between firms from ERC and SEC. Lastly, accounting comparability decreases stronger between firms from ERC and SEC than between firms from SEC and WEC but with a statistically smaller level of significance.

Comparing the levels of accounting comparability in the pre-adoption period with the post-adoption period, they are significantly lower for all three cluster pairs in the post-adoption period. From the first to the third period, accounting comparability decreased by 0.1021 between firms from ERC and SEC, by 0.0360 between firms from ERC and WEC as well as by 0.0390 between firms from SEC and WEC. Obviously, the decrease has been strongest between firms from ERC and SEC. Between firms from ERC and WEC and firms from SEC and WEC the magnitude of the decrease did not differ significantly.

Summing up the results of the first accounting comparability metric (metric 3.1), accounting comparability increases significantly for all three cluster pairs between the pre-adoption period and the adoption period. During the post-adoption period it decreases again for all three cluster pairs, reaching significantly lower levels of accounting comparability than in the pre-adoption period. Hence, over the whole period covered by this thesis, there was a decrease in accounting comparability according to this metric. Interestingly, there have been significant differences in accounting comparability between different cluster pairs initially. Over time, however, there is a convergence of accounting comparability between all three cluster pairs.

6.2.1.2 *AGG.CFCOMP* metric 3.2

Metric 3.2, the cash flow-based accounting comparability metric, *AGG.CFCOMP*, yields quite contrasting results (see *Table 14* and *Figure 5*). According to this metric, accounting comparability between firms from ERC and WEC (-0.2322) does not differ significantly from accounting comparability between firms from SEC and WEC (-0.2278) in the pre-adoption period. The accounting comparability between firms from ERC and SEC (-0.3109) is significantly lower compared to both other cluster pairs. In the adoption period accounting comparability between firms from ERC and SEC (-0.8864) again is significantly lower compared to both other cluster pairs but this time, also accounting comparability between firms from ERC and WEC (-0.7496) is significantly lower than between firms from SEC and WEC (-0.3224).

The results indicate that accounting comparability decreased significantly for all three cluster pairs from the first to the second period: between firms from ERC and SEC by 0.5754, between firms from ERC and WEC by 0.5175 as well as between firms from SEC and WEC by 0.0946.

These results are completely opposite to the findings of metric 3.1 which indicates an increase in accounting comparability for all cluster pairs in the adoption period. For metric 3.2, the decrease in accounting comparability has been significantly stronger for accounting comparability between firms from ERC and SEC compared to the other two cluster pairs. Also, the decrease is significantly stronger for accounting comparability between firms from ERC and WEC than between firms from SEC and WEC.

			Cluster pair(s)						
			Differences						
			ERC vs. SEC	ERC vs. WEC	SEC vs. WEC	(ERC vs. SEC) - (ERC vs. WEC)	(ERC vs. SEC) - (SEC vs. WEC)	(ERC vs. WEC) - (SEC vs. WEC)	
Period(s)	Differences	1	-0.3109	-0.2322	-0.2278	-0.0787***	-0.0831***	-0.0043***	
		2	-0.8864	-0.7496	-0.3224	-0.1367***	-0.5640***	-0.4272***	
		3	-0.4654	-0.2002	-0.4053	-0.2652***	-0.0601***	0.2051***	
		(1)-(2)	0.5754***	0.5175***	0.0946***	0.058***	0.4809***	0.4229***	Difference-in-differences
		(1)-(3)	0.1545***	-0.0320***	0.1774***	0.1864***	-0.023***	-0.2094***	
		(2)-(3)	-0.4210***	-0.5494***	0.0829***	0.1285***	-0.5038***	-0.6323***	
						Differences-in-differences			

The table presents the results of the second, cash flow-based accounting comparability metric from equation 24 for the aggregated average accounting comparability between two clusters for each period (upper-left segment), the differences between two cluster pairs for distinct periods (upper-right segment), the differences between periods for distinct cluster pairs (lower-left segment) and the differences-in-differences between periods and cluster pairs (lower-right segment). The differences-in-differences compare the differences between periods over cluster pairs. We number the periods for ease of presentation where period 1 is the pre-adoption period, period 2 the adoption period and finally, period 3 the post-adoption period. Values are rounded to four decimal places. We test significance using a percentile-bootstrapping procedure. Thus, we get the empirical distribution of the differences of the metric. *, ** and *** indicate significance at the

Table 14: Metric 3.2 *AGG.CFCOMP* (Panel 7)

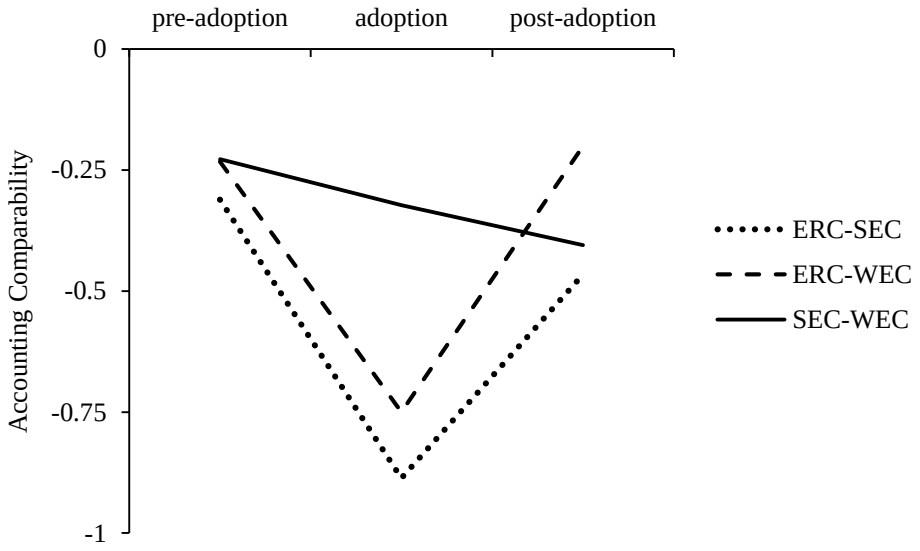

Figure 5: Cash flow-based accounting comparability metric

In the post-adoption period, the relative ranking of the cluster pairs in terms of accounting comparability changes: Now, accounting comparability is significantly higher between firms from ERC and WEC (-0.2002) compared to the other two cluster pairs. The comparability between firms from SEC and WEC (-0.4053) also is significantly higher than accounting

comparability between firms from ERC and SEC (-0.4654). In contrast to metric 3.1, there is no convergence of accounting comparability between the three cluster pairs in the post-adoption period. Regarding the change in accounting comparability from the second to the third period, accounting comparability increases between firms from ERC and SEC by -0.4210 and between firms from ERC and WEC by -0.5494, but decreases further between firms from SEC and WEC by 0.0829. This is interesting because it is the only time for both metrics that accounting comparability did not change into the same direction for all cluster pairs in one period.

Furthermore, the increase in accounting comparability is significantly stronger between firms from ERC and WEC than between firms from ERC and SEC. Comparing the levels of accounting comparability in the pre-adoption period with the post-adoption period, accounting comparability between firms from ERC and SEC decreases significantly by 0.1545 from the first to the third period. Likewise, accounting comparability between firms from SEC and WEC decreases significantly by 0.1774. Only between firms from ERC and WEC accounting comparability increases from the first to the third period by -0.0320. Obviously, the change is significantly stronger for accounting comparability between firms from ERC and SEC as well as between firms from SEC and WEC than between firms from ERC and WEC. However, the development of accounting comparability between firms from SEC and WEC is much more stable and less erratic as it decreases slightly and, most of all, steadily. Without the very strong decrease in accounting comparability between firms from ERC and the two other clusters in the second period, the development of accounting comparability between all three cluster pairs would have been much more similar. Hence, certain developments in ERC in the second period must have had a strong impact on the general picture of the results of this metric. We come back to this issue in our discussion (*Section* 7).

Summing up, the results of the cash flow based accounting comparability metric (metric 3.2) are in strong contrast to the findings of the return based accounting comparability metric (metric 3.1). The cash flow based metric shows a decrease of accounting comparability for all cluster pairs in the adoption period. In the post-adoption period, it indicates that accounting comparability increases again for two cluster pairs and only decreases further for one cluster pair. Over the three periods, accounting comparability decreases for two cluster pairs and increases slightly for one cluster pair. Consequently, the cash flow-based metric (metric 3.2) shows no convergence of accounting comparability between all three cluster pairs as indicated by metric 3.1 but rather a continuing divergence in accounting comparability between the three cluster pairs.

6.2.2 Evaluation of hypotheses

Based on the empirical results, we next evaluate our hypotheses concerning accounting comparability (H4 to H6). Based on the findings of prior literature, we expected in H4 that accounting comparability between firms from ERC and SEC will increase after IFRS adoption because the enforcement reforms in ERC lead to a convergence of reporting incentives and hence a similar interpretation of discretionary accounting choices by firms in the two clusters. Further, we hypothesized in H5 and H6 that accounting comparability between firms from ERC and WEC as well as between firms from SEC and WEC will show no increase in the long run because the incentives to report high quality financial statements remain low for firms from

WEC also after IFRS adoption. The opportunistic use of accounting discretion by firms from WEC was argued to prevent an increase in accounting comparability for these two cluster pairs. *Figure 4* and *Figure 5* illustrate that the two accounting comparability metrics yield almost diametrically opposing results. As there are only two metrics it is not possible to consider what the majority of metrics indicates and derive a general trend as we did for the accounting quality metrics. Nevertheless, H4 can be rejected as neither of the two metrics shows an overall increase in accounting comparability between firms from ERC and SEC from the pre-adoption period to the post-adoption period. H5 can neither be rejected nor confirmed because the return-based metric indicates a decrease in accounting comparability between firms from ERC and WEC in the long run, but the cash flow-based metric shows an increase in accounting comparability. H6 can be confirmed in that accounting comparability between firms from SEC and WEC did not increase in the long run. However, the decrease for this cluster pair comes somewhat surprising as we rather expected a more or less stable development of accounting comparability. For this reason, we discuss the implications of our results, e.g. concerning the role of enforcement, as well as the contrasting results of the two metrics in more detail in the following section.

7 Discussion

Having identified certain patterns in the empirical results, in this section the results both of accounting quality and comparability are interpreted and compared to the findings of prior research in order to identify similarities and discrepancies. Especially, we discuss the role of IFRS and enforcement on the first-order effects of IFRS adoption as well as the implications of our results for the existence of potential temporary learning effects.

7.1 Accounting Quality

7.1.1 Comparison with prior research

Comparing our results for accounting quality with prior research, we find that they are in line with the results of Ahmed *et al.* (2012) who also report a decrease in most accounting quality measures for IFRS adopters relative to a benchmark sample of non-adopters thus showing that the results do not simply stem from a general decline in accounting quality. Notwithstanding, as they only analyze a two-year period after IFRS adoption it did not become clear whether the observed effects are of permanent or temporary nature. As our results show a reduction in accounting quality over a seven-year period after IFRS adoption, they indicate that there is rather a permanent decrease in accounting quality of firms from IFRS adopting countries. Nevertheless, potential learning effects need to be addressed in this context and will therefore be discussed later. Additionally, Ahmed *et al.* (2012) find that their results are primarily driven by strong enforcement countries. This seems to be contradicting to what we find as we observe a stronger decrease for weak enforcement countries. However, their sample clustering differentiates between weak and strong enforcement countries solely based on the Rule of Law index of Kaufmann *et al.* (2009). This index reflects general differences in the strength of the legal system (such as the quality of the police, courts as well as the likelihood of crime) rather than specific differences in enforcement systems. On the contrary, the clustering of our sample

countries is based on six enforcement variables from La Porta *et al.* (2006) and Djankov *et al.* (2008) which, in our view, is therefore better suited to distinguish country clusters according to their differences in enforcement. Unsurprisingly, the resulting composition of country clusters in Ahmed *et al.* (2012) and our thesis differs so that the seeming contradiction about the strongest drivers behind the decrease in accounting quality on a closer look disappears.

However when comparing our results with the findings of Barth *et al.* (2008), there are considerable differences as Barth *et al.* (2008) report an increase of accounting quality after IAS adoption. Still, these differences can be potentially explained for because they analyze a sample of voluntary IAS adopters in the period from 1994 to 2003. Hence, first their results might be driven by a self-selection bias due to the sole focus on voluntary adopters whereas our thesis includes both mandatory and voluntary adopters. Second, they analyze a time span almost one decade before our sample period. Hence, differences in general trends and economic characteristics might drive the differences between the findings of their study and our thesis. Last but not least, Barth *et al.* (2008) analyze a sample comprising only 1,896 firm-year observations whereas our results are based on a much larger sample size (depending on the test ranging from 40,000 to 60,000 firm-year observations).

As discussed before, Samarasekera *et al.* (2012) find an increase in accounting quality of UK-firms subject to strong enforcement mechanisms. Thus, their findings seem contradicting because our findings do not indicate increases in accounting quality neither for strong enforcement countries nor for countries which implemented concurrent enforcement reforms. However, this might be due to the focus of Samarasekera *et al.* (2012) on a single country, namely the UK. Moreover, it might be the case that accounting quality indeed increased in the UK but decreased in the other countries of the ERC cluster so that this effect does not become visible in our aggregated cluster results (we will discuss the special case of the UK in a robustness test later in *Section 7.3.1*). More probably, though, the difference might stem from a different level of analysis. While we differentiate countries according to macro institutional enforcement reforms (e.g. the introduction of a proactive review process by enforcement agencies) and the set of six enforcement variables (see *Appendix 3*), Samarasekera *et al.* (2012) argue that cross-listing of firms increases enforcement for these firms as they become subject to extended financial reporting requirements and enhanced auditor oversight. Thus while we treat all firms of a country equally and focus on macro country differences, Samarasekera *et al.* (2012) are rather interested in the reporting incentives on the firm-level. Again, it might be the case that accounting quality increases for a subset of firms of a specific country cluster but when looking at the results of the cluster an overall decrease in accounting quality can be observed. Summing up, the role of enforcement on accounting quality can still not be finally determined and requires further discussion.

7.1.2 The role of enforcement for accounting quality

Our findings show that accounting quality differs across most country clusters of our sample and as these clusters differ with respect to enforcement characteristics, the differences in accounting quality between the clusters indicate that enforcement does have an influence on accounting quality. This is in line with theory (Soderstrom and Sun, 2007) and prior empirical studies (e.g. Christensen *et al.*, 2012; Samarasekera *et al.*, 2012) which find that enforcement

plays an important role for explaining the effects of IFRS adoption. However, looking at the firms from ERC we find a decrease of accounting quality even though these countries implemented enforcement reforms. At this point, one could argue that the reforms at least mitigated the decrease in accounting quality because the decrease for firms from WEC – which did not realize enforcement reforms – was more pronounced from the pre-adoption to the adoption period than for firms from ERC. Nevertheless, one needs to keep in mind that the differences-in-differences (i.e. the difference in the magnitude of the change in accounting quality) between ERC and WEC from the first to the second period show no consistent picture. While the results consistently show that accounting quality decreased stronger for firms from WEC than for firms from SEC, only the correlation between cash flow and accrual metric (metric 1.3) indicates that accounting quality decreased stronger for WEC than for ERC. The other two earnings smoothing metrics (metrics 1.1 and 1.2) indicate that accounting quality actually decreased even stronger for ERC than for WEC, despite the reforms undertaken. In other words, a mitigating effect of enforcement reforms cannot be detected.

Interestingly, this result is in strong contrast to the findings of Christensen *et al.* (2012) who find that second-order market liquidity effects following IFRS adoption were restricted to ERC. Thus, while enforcement reforms seem to have influence on capital market effects (second-order effects) they do not actually change the average quality of the underlying financial statements (first-order effect). A possible explanation for this discrepancy might be that the enforcement reforms realized in some countries seem to have created positive effects on the trust of investors in financial statements, thus leading to the observed positive capital market effects, without changing the actual quality underlying financial statements.

Another argument against a strong influence of enforcement reforms on accounting quality is that the cluster differences in accounting quality already exist in the pre-adoption period and that the relative ranking of the clusters remains stable over time even though some countries implemented significant changes in their enforcement systems. If enforcement reforms played such a big role, then these changes should have affected the relative ranking of the clusters in terms of accounting quality in the periods following the enforcement reforms (i.e. firms from ERC should have shown similar accounting quality compared to firms from SEC only after the enforcement reforms and not already in the pre-adoption period). However, if the enforcement reforms in the ERC cluster which were quite substantial do not have major consequences on accounting quality, then the role of enforcement itself on accounting quality is at least questionable. To sum up this central finding of our thesis, our results indicate that enforcement, as part of a country's institutional framework, has certain effects on accounting quality but does not play such a dominant role when explaining accounting quality effects of IFRS adoption as suggested by other studies (Ahmed *et al.*, 2012; Samarasekera *et al.*, 2012).

7.1.3 Further implications

A further question is whether our thesis provides insights in other factors potentially influencing accounting quality. Concerning the general discussion whether IFRS adoption led to higher accounting quality or not, and thus reached one of its main objectives, our thesis indicates that this is not the case as accounting quality decreases after IFRS adoption. However, as prior research (e.g. Soderstrom and Sun, 2007) shows that apart from the financial reporting

standards several other factors influence accounting quality, it cannot be concluded that IFRS adoption is responsible for the witnessed decrease of accounting quality. Nevertheless, the findings indicate that IFRS adoption did not lead to an increase in accounting quality. This is in line with prior research suggesting that reporting incentives dominate accounting standards with regard to the influence on first- and second-order effects (Christensen *et al.*, 2008).

Consequently, there might be other factors influencing the differences in accounting quality across clusters if both IFRS and enforcement do not seem to be the main drivers. In this context, it is important that the differences in accounting quality between the clusters remain stable over all three periods. In their study, Barth *et al.* (2008) find that in the pre-adoption period there are no significant differences in accounting quality between their two subsamples (IAS adopters and non-adopters). Hence, they trace the differences in accounting quality in the post-adoption period back to IFRS adoption and rule out that differences in economic characteristics are the drivers of these differences (because in that case they should already have existed in the pre-adoption period). In our case, differences in accounting quality already exist before both IFRS adoption and the implementation of the enforcement reforms in ERC. While we argued that the clustering of countries is based on differences in enforcement systems, it might be the case that our country clusters also differ in other institutional and economic characteristics and that these differences are driving the differences in accounting quality. These differences could then already be noticed in the first period. This argument is supported by the existence of several significant differences-in-differences over all measures because – as explained before – significant differences-in-differences imply that accounting quality changes with different magnitudes for the clusters. Accordingly, the clusters are subject to different underlying trends. However, what exactly constitutes these potentially institutional or economic characteristics are, is a question that cannot be addressed within the scope of our thesis.

Another area for future research might be the topic of potential learning effects following IFRS adoption. In *Section 3.1*, we discussed that following IFRS adoption accounting quality might fluctuate because preparers of financial statements (i.e. firms and auditors) need some time to adjust and adapt their reporting practices to the new financial reporting standards. Likewise, following the enforcement reforms in some countries the enforcement authorities might need some time to get used to and effectively enforce the new rules. As a result, we argued that during the adoption period there might be temporary changes in accounting quality and comparability. It is for exactly this reason that we use three instead of two periods in our analysis. Looking at our results, however, it is not possible to finally answer this question. The results from the post-adoption period generally indicate that accounting quality did not change significantly compared to the adoption period. Nonetheless, some metrics suggest an increase (metric 1.1 for ERC and WEC) or decrease (metrics 1.4 for ERC and WEC; 1.1, 1.3, and 1.5 for SEC) of accounting quality of single clusters.

Assuming that accounting quality actually increases again between the second and the third period, this would imply an upward trend of accounting quality even though on an absolute level still being lower than in the pre-adoption period. The decline of accounting quality in the adoption period might then be traced back to temporary adjustment problems which are overcome by learning effects leading to an increase in accounting quality in the post-adoption

period. Leuz (2010) points out that institutional settings are 'sticky' and change only slowly over time. Hence, it might be the case that learning and adjustment processes take longer than expected and the full effects of IFRS adoption in combination with changes in enforcement regimes emerge only after considerable time. For this reason, future research covering longer time periods might shed additional light on this issue. Nevertheless, the scenario described is only one possible scenario. It might equally be the case that accounting quality remains stable or decreases even further in the post-adoption period. In that case, it could be ruled out that decreases in accounting quality are only due to adoption and learning effects. This question, however, cannot be answered at this point in time.

Leuz (2010) also argues that there are robust country clusters of institutional differences. As there are considerable interdependencies between single elements of the institutional framework (including accounting standards), he argues that despite the harmonization of accounting standards, differences in reporting practices will exist and persist due to the stickiness of institutional settings. Our findings confirm the existence of such a continuing divergence of reporting practices. Indeed there are differences in accounting quality between the clusters and they continue to exist also in the post-adoption period. In other words, reporting practices do not converge as this would imply that there are no significant differences between the clusters in the third period any longer. Therefore, another finding of our thesis is that despite the harmonization of accounting standards reporting practices continue to differ between countries.

Finally, the value relevance metrics do not allow to draw significant conclusions. This is more or less in line with the findings of Barth *et al.* (2008). They find that value relevance for IFRS adopters as a joint group does not significantly differ from non-adopters in the pre-adoption period. In their post-adoption period, only the first value relevant metric (metric 2.1) indicates that value relevance increases for IFRS-adopters. The other two metrics do not show significant differences between the two clusters of their study. Our thesis extends these findings by showing that there neither are significant differences concerning value relevance within subclusters of the IFRS adopting countries. Samarasekera *et al.* (2012), however, do find an increase of value relevance for IFRS adopters (for both cross-listed and not cross-listed firms). Nevertheless, the results of their study again cannot be compared without restrictions to our findings because they only analyze a single-country setting. Also, in additional tests the authors are not able to identify the reasons for the improvement in value relevance under IFRS compared to UK-GAAP. Hence, it is not possible to fully align the findings on value relevance of our thesis with findings of prior studies.

Apart from that, the results on value relevance have to be interpreted cautiously. While the analyses of earnings management allow inferences on the development of accounting quality, the three value relevance metrics sometimes even yield diametrically opposing results. One reason for these differences might be that not all value relevance metrics have been controlled sufficiently for mean differences between countries and industries: For the calculation of metric 2.1, in the first step the stock price, *P*, has been regressed on country and industry fixed effects. However, net income per share, *NIPS*, has not been controlled for such fixed effects. Hence, the association between stock prices and earnings might at least partly be distorted. A

similar problem relates to metrics 2.2 and 2.3. Here, net income scaled by stock price, *NI/P*, has been controlled for country and industry fixed effects, but not annual stock return, *RETURN*. However, if there are differences in market efficiency between countries, then returns are affected by country fixed effects. Not controlling for them might lead to a distortion of the results. Furthermore, Runsten (1998) finds that value relevance is related to the business cycle, indicating that value relevance is lower in times of economic boom and higher during economic recessions. Hence, for all metrics both the independent and dependent variables might need to be controlled for further influencing factors such as economic growth rates. For these reasons, concerns must be raised whether the applied metrics are good measures of accounting quality. We chose them in order to be in line with the metrics applied by prior research. However, improving the value relevance metrics applied in studies on accounting quality is another interesting area for future research.

7.2 Accounting Comparability

7.2.1 Discrepancy between the two metrics

Concerning accounting comparability, the empirical results provide no clear picture as the two metrics yield very dissimilar results. However, this tension between the two metrics is very interesting because it explains much about what the metrics actually capture. As both metrics have been introduced only recently and therefore not been applied extensively, our results help to evaluate their usefulness in describing accounting comparability. Hence, before being able to draw conclusions about the role of enforcement on accounting comparability or the existence of potential learning effects, the opposing results of these two metrics need to be analyzed. In this context, it is necessary to understand the differences in the nature of the two metrics.

Cascino and Gassen (2012) point out that the return-based accounting comparability metric (metric 3.1) as suggested by De Franco *et al.* (2011) might be influenced by differences in market efficiency between different countries as well as other market imperfections as these factors affect stock returns. For this reason, Cascino and Gassen (2012) introduce the second cash flow-based accounting comparability metric (metric 3.2). The idea of De Franco *et al.* (2011) is that accounting comparability can be analyzed by comparing the mapping of economic events into financial statements of different firms. Accordingly, both metrics capture different mappings because they use different proxies for economic events. The return-based metric captures the mapping of stock returns into earnings whereas the cash flow-based metric captures the mapping of cash flows into earnings. If both metrics yield different results, these differences must be due to factors which influence the mapping of cash flows into earnings in a different way than the mapping of stock returns into earnings. In the following, we present potential explanations for the observed differences in the results.

In the adoption period, the increase in accounting comparability for all cluster pairs indicated by the return-based metric (metric 3.1) might be driven by co-movements of stock returns due to the financial crisis. More specific, the economic recession might have decreased both stock returns and earnings of firms from different clusters in a similar way. Thus, the mapping of returns into earnings would have become more similar between the clusters indicating an increase in accounting comparability, even though there are not necessarily changes in the way

preparers of financial statements use discretionary accounting choices. Hence, the increase in accounting comparability could be described as a *pseudo-increase* due to a convergence of the correlation between returns and earnings between clusters but no real convergence of reporting practices. This notion is supported by two other aspects: the fact that the cash flow-based metric does not indicate an increase in accounting comparability during that period and the fact that also the return-based metric shows a decrease in accounting comparability in the post-adoption period when co-movements of earnings and stock returns decreased again compared to the years of the financial crisis. This implies that the return-based metric might not be optimal for measuring accounting comparability across countries because capital market effects and differences in market efficiency which are unrelated to actual reporting practices seem to have a strong influence on the metric (Ball and Shivakumar, 2006; Holthausen, 2003).

As a next step, we discuss potential explanations of the strong decrease in accounting comparability, especially for the cluster pairs including ERC, during the adoption period indicated by the cash flow-based metric. If the cash flow-based metric indicates a decrease in accounting comparability but the return-based metric does not, then the decrease in the cash flow-based metric must stem from factors that only have an effect on the correlation between cash flows and earnings but not on the correlation between returns and earnings. This could be, for instance, non-cash charges such as impairment. If firms from different clusters handled the new impairment rules under IFRS in different ways, then even though their cash flows would not be affected by their impairment policy, their earnings would be. Consequently, the mapping of cash flows into earnings would differ between the clusters and the cash flow-based accounting comparability metric would report a decrease in accounting comparability. This notion is supported by the debate on IAS 36 which shows that in particular immediately after IFRS adoption there might indeed have been differences in the interpretation of the standard across our clusters (Nobes, 2006). This would in turn imply that there actually was a decrease in accounting comparability during the adoption period. Summing up so far, the cash flow-based accounting comparability metric seems to be more adequate for measuring accounting comparability, especially in a multi-country analysis, because it is not affected by capital market effects or differences in market efficiency.

In addition, this discussion shows that interpretations of the two accounting comparability metrics need to be done carefully. Although certain effects can be observed, concrete inferences about drivers of the effects are hardly possible due to the high aggregation level of the metrics. Clearly, this is another point of critique of the two metrics but one also needs to acknowledge that e.g. the discussion of the accounting quality results also shows how difficult it is to isolate single effects in such complex and multi-faceted research settings.

7.2.2 Further implications

As the results of the return-based accounting comparability metric (metric 3.1) need to be interpreted very carefully, the following comparison of our results with prior research and the further discussion focuses on the results of the cash flow-based accounting comparability metric (metric 3.2).

Comparing our findings with prior research, Cascino and Gassen (2012) find that accounting comparability only marginally increases after IFRS adoption when controlling for firm size and audit quality. This implies that the adoption of IFRS alone has only limited effects on accounting comparability and shows that reporting incentives dominate accounting standards. Our thesis, however, goes a step further by differentiating different groups of IFRS adopting countries and analyzing whether the groups experience different developments of accounting comparability. Nevertheless, it is very difficult to draw conclusions from our results on whether IFRS adoption led to an increase in accounting comparability or not. For two cluster pairs, accounting comparability decreases overall from the pre-adoption to the post-adoption period and for the remaining cluster pair there is only a small increase in accounting comparability. Therefore, our findings can be interpreted as further evidence that IFRS did not lead to a substantial increase in accounting comparability. Hence, our thesis suggests that the introduction of IFRS might have failed to reach one of its major objectives. This is in line with Lang *et al.* (2010) who also find a significant decrease in accounting comparability measured with the methodology of De Franco *et al.* (2011) following IFRS adoption. However, Lang *et al.* (2010) do not differentiate between the institutional differences of their sample and therefore their results and ours cannot be compared in detail.

Trying to evaluate the role of enforcement for accounting comparability, it can be noticed that between the three cluster pairs there mostly are significant differences in accounting comparability measured by the cash flow-based metric. As the clustering was performed according to differences in countries' enforcement regimes, it can be inferred that enforcement seems to play a role for accounting comparability. Nevertheless, due to the high aggregation of the metrics, further implications would belong to the realm of speculation. For example, one can only speculate why firms from ERC might have handled impairment differently in the adoption period so that accounting comparability between firms from ERC and the other two clusters decreased so strongly. Hence, no inference about the role of enforcement reforms for accounting comparability can be made based on our findings.

Similarly, no reliable conclusions can be drawn concerning the existence of potential learning effects during the adoption period. While this is already difficult in the context of accounting quality, at this point we can only state that the increase of accounting comparability of firms from ERC with firms from the other two clusters in the post-adoption period might indicate that firms from ERC needed some time to adjust to the new accounting standards. Regarding the question posed by Leuz (2010) whether there is a divergence or convergence of reporting practices between clusters following IFRS adoption, our results can be interpreted more clearly: the cash flow-based accounting comparability metric shows significant differences in accounting comparability for all three cluster pairs, also in the post-adoption period, which is additional evidence of a continuing divergence of reporting practices.

Finally, it makes sense to combine the results of accounting quality and comparability. Our research design addressed both first-order effects simultaneously because increases in accounting comparability are worth little if they come along with decreases in accounting quality. From a theoretical point of view, a decrease in accounting quality of only one cluster would lead to a decrease in accounting comparability because the increasing opportunistic use

of discretionary accounting choices by firms from one cluster would reduce accounting comparability with firms from another cluster. Still, high accounting quality is no pre-condition for high accounting comparability because if accounting quality decreases evenly for different clusters (e.g. if the level of earnings smoothing increases equally), then the mapping of cash flows into (then smoother) earnings changes equally for all clusters, leading to an increase of accounting comparability. Our results generally show a decrease in accounting quality from the pre-adoption to the post-adoption period and a simultaneous decrease in accounting comparability (measured by the cash flow-based metric) for most cluster pairs. As argued, these findings make sense from a theoretical point of view.

7.3 Robustness tests

Having discussed both the results on accounting quality and comparability, in the following, we perform robustness tests to control for potential shortcomings in our research design.

7.3.1 The case of the UK

As discussed in *Section 4.2*, an alternative clustering method would have been to also include ERC into the k-means cluster analysis instead of ex-ante assigning them to the first cluster. The alternative method results in the same country clusters except for the case of the UK which then belongs to the SEC cluster and not the cluster comprising Finland, Germany, the Netherlands and Norway. This outcome is quite in line with intuition as the UK also represents an Anglo-Saxon country. Because of the different results of the two clustering techniques we performed a robustness test by leaving the UK out of the sample to see whether the findings are driven by the UK. An alternative of how to perform this robustness test would be not to leave out the UK completely, but rather assign it to the SEC cluster. However, thereby two clusters would be affected making changes in the results less easy to interpret. Moreover, unlike the other SEC the UK did perform concurrent enforcement reforms wherefore putting them together would decrease the homogeneity within the cluster.

7.3.1.1 Accounting quality

The results of the robustness test concerning accounting quality again only allow conclusions about earnings management and not about value relevance.[15] The tenor of the inferences is not changed. *Figure 6* summarizes the development of accounting quality in terms of earnings management leaving the UK out of the sample. Unsurprisingly, the results concerning firms from WEC and SEC are the same because these clusters have not been changed.[16] Concerning ERC, there also is a decrease in accounting quality from the pre-adoption to the adoption period. For all clusters, it again is not possible to pinpoint the exact development of accounting quality in the post-adoption period as the metrics show ambiguous developments. For ease of illustration, in *Figure 6* the two ranges of potential outcomes of accounting quality in the post-adoption period for firms from ERC and WEC have been combined (as otherwise there would be two overlapping shaded areas). The only major change between the results including the UK

[15] The results of the five earnings management and the comparability metrics are tabulated in *Appendix 8*. The results on value relevance are not tabulated because they do not allow any important inferences.

[16] Note that the values of the accounting quality metrics, however, differ very slightly from the values including the UK due to the winsorizing.

and the results excluding the UK lies in the relative ranking of the clusters in terms of accounting quality. While the results including the UK (see *Figure 3*) show no significant differences between firms from ERC and SEC for all three periods, the results excluding the UK only do so in the pre-adoption period. In the adoption and post-adoption period firms from SEC now exhibit a significantly higher accounting quality than firms from ERC. Comparing ERC and WEC, the findings including the UK indicate that firms from ERC have a higher accounting quality in all three periods. The results from the robustness test, however, do so only for the pre-adoption period. In the adoption and post-adoption period accounting quality of firms from ERC has declined from the level of firms from SEC to the level of firms from WEC, showing no significant differences compared to firms from WEC anymore. Summing up, the results of the robustness test excluding the UK indicate that the UK plays an important role for the ERC cluster because without the UK accounting quality of firms from ERC is significantly lower (at the same level as of firms from WEC instead of SEC).

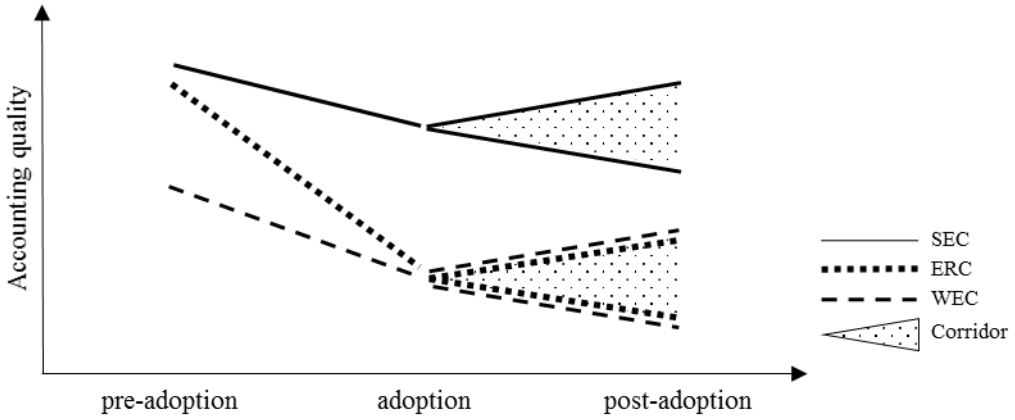

Figure 6: Accounting quality in terms of earnings management excluding the UK

However, while leaving the UK out of the sample alters the relative ranking of the clusters in terms of accounting quality, the inferences from the results mostly remain unchanged. On the contrary, the importance of enforcement reforms for accounting quality seems to be even lower when taking the UK out of the sample. Previously, firms from ERC at least had an accounting quality similar to firms from SEC. Without the UK, accounting quality of firms from ERC decreases even stronger in the adoption period until there are no significant differences even to firms from WEC any more, even though substantial enforcement reforms have been implemented also in the remaining countries of the ERC cluster. This implies that the effect of enforcement reforms on accounting quality is marginal at best. Concerning the other inferences, there still is a continuing divergence of reporting practices, also when leaving the UK out of the sample because differences in accounting quality continue to exist in the post-adoption period. Finally, still no final answer can be given concerning the existence of potential learning effects in the adoption period because no unambiguous conclusions can be drawn about the development of accounting quality in the last period.

7.3.1.2 Accounting comparability

Similarly, we re-calculate the results of the two accounting comparability metrics without including the UK into the sample.[17] *Figure 7* and *Figure 8* show the results of the two metrics of the robustness test excluding the UK. Necessarily, the results for accounting comparability between firms from SEC and WEC are exactly the same. Accounting comparability between firms from the other two cluster pairs does change a bit. As we do not test for significance we cannot evaluate whether the changes have been significant. However, when comparing the graphs of the robustness test (*Figure 7* and *Figure 8*) with the results including the UK (*Figure* 4 and *Figure 5*), no big differences can be noticed. For instance, in *Figure 8* the line representing accounting comparability between firms from ERC and WEC moves slightly up compared to *Figure 5*, indicating that without the UK accounting comparability between firms from the two clusters is a bit higher. The relative ranking between the three cluster pairs also changes slightly in some cases. Importantly, the change in accounting comparability of the three cluster pairs between the periods is not affected at all (e.g. no increases instead of decreases).

Summing up, the UK also slightly changes the results on accounting comparability of firms from ERC with firms from the other two clusters. However, no major changes occur. As it is not possible to draw detailed conclusions from the results of the two metrics due to their high aggregation, our inferences about accounting comparability are not affected by the UK.

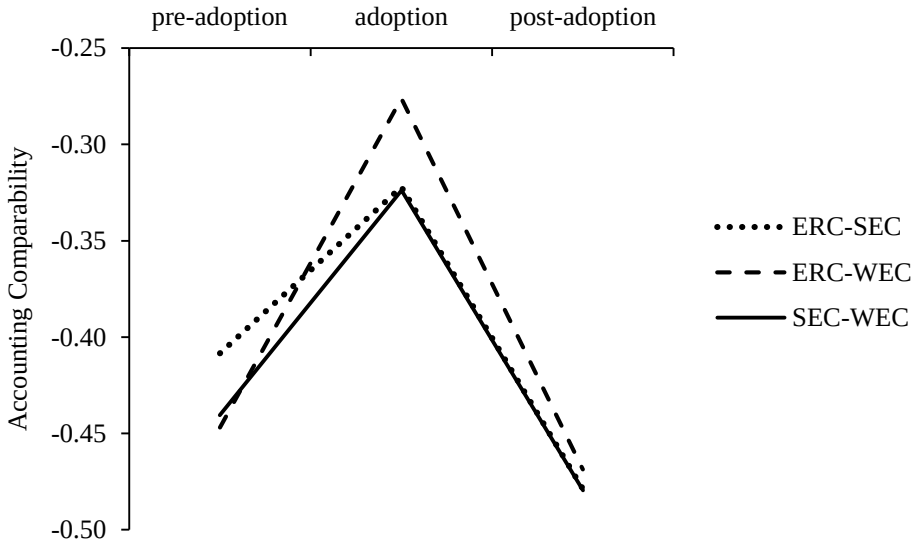

Figure 7: Return-based accounting comparability metric excluding the UK

[17] However, we do not test for significant differences between the results of each cluster pair because a bootstrapping procedure with 30 replications would have taken over two weeks.

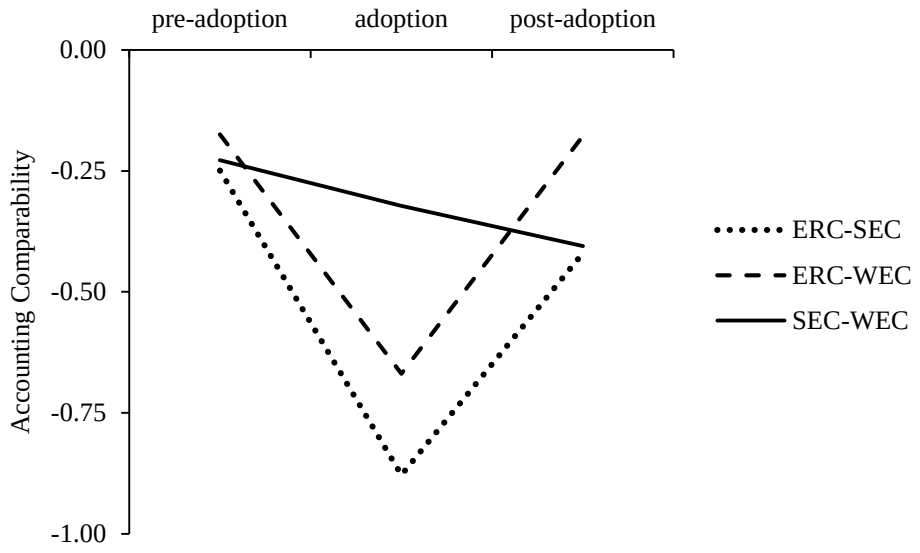

Figure 8: Cash flow-based accounting comparability metric excluding the UK

7.3.2 Subsequent enforcement reforms

Another issue that needs to be analyzed arises because three countries implemented substantial reforms of their enforcement regimes, but not at the same point in time as countries from the ERC cluster, but subsequent to IFRS adoption. While Sweden (SE) realized such reforms in 2007, Hong Kong (HK) and Turkey did in 2008. Assigning these countries to a separate fourth cluster is not practical because the complexity and duration of calculations would have increased exponentially. However, it might be the case that the enforcement reforms in these countries led to a different development of accounting quality in the post-adoption period compared to the other countries of the their respective cluster. As Hong Kong for instance comprises roughly one third of the firm-year observations of the SEC cluster (see *Table 4*), results might be driven by these countries. To find out whether these three countries have a strong effect on the general picture of our results, another robustness test is performed leaving the three countries out of the sample.[18]

The results of the robustness test only allow conclusions about earnings management but not about value relevance.[19] *Figure 9* summarizes the development of accounting quality in terms of earnings management leaving Sweden, Hong Kong and Turkey out of the sample. The results indicate that the three countries do not change the general picture of our findings. Accordingly, *Figure 9* illustrating the results without the three countries does not differ from *Figure 3* illustrating the results including the three countries in any aspect. Hence, the different timing of the enforcement reforms in Sweden, Hong Kong and Turkey does not change the inferences from our results. Given that the discussion of our results suggested that the role of enforcement reforms on accounting quality is rather marginal, the findings of the robustness test are not surprising.

[18] In this test, we only re-calculate the accounting quality metrics 1.1-2.3 because the calculation of the accounting comparability metrics 3.1 and 3.2 takes such a long time.

[19] The results of the five earnings management metrics are tabulated in Appendix 10. The results on value relevance are not tabulated because the results do not allow any important inferences.

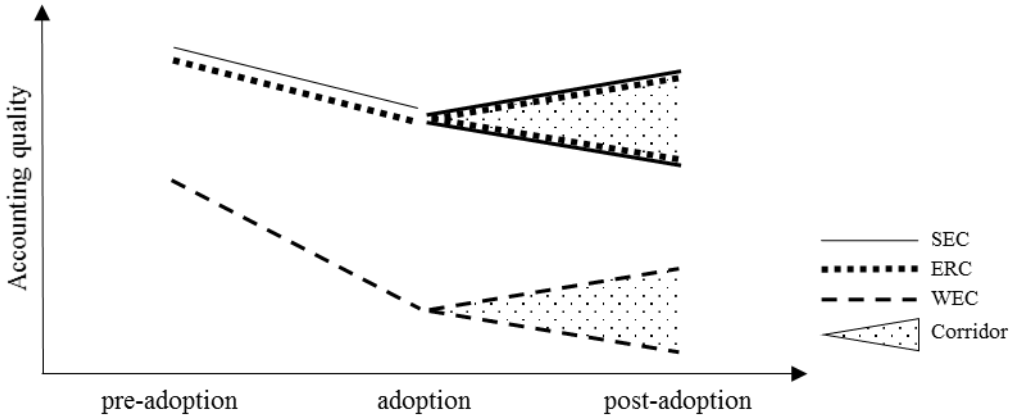

Figure 9: Accounting quality excluding HK, SE, TR

7.3.3 Different IFRS adoption dates

Finally, another potential shortcoming in our research design comes from the fact that so far we did not adjust the time periods for countries with IFRS adoption dates differing from the end of 2005. However, Singapore (SG) adopted IFRS in 2003, Turkey in 2006, New Zealand (NZ) and Pakistan (PK) in 2007, and Israel in 2008. Thus, our results might be distorted by the fact that some effects from this group of countries might be observed in a different time period blurring the overall results. Originally, we did not adjust for this circumstance because when introducing relative time periods (e.g. for SG a pre-adoption period from 1999-2002, an adoption period from 2003-2006 and so on) all other effects, such as macroeconomic effects, are not captured correctly anymore and may even cause stronger distortions. For instance, the years of the financial crisis would then be assigned to different periods which – assuming the crisis has an effect on accounting quality and comparability – might lead to a distortion of the results. However, a robustness check leaving out these five countries is necessary to investigate whether these differences of time period have a profound impact on the results. An analysis without the five countries therefore leads to a better identification of the IFRS' effects on accounting quality and comparability.

7.3.3.1 Accounting quality

The results of the robustness tests concerning accounting quality again only allow conclusions about earnings management and not about value relevance.[20] *Figure 10* summarizes the development of accounting quality in terms of earnings management leaving Singapore, Israel, New Zealand, Pakistan and Turkey out of the sample. The similarity to *Figure 3* indicates that the five countries do not change the tenor of our inferences. Again, we find that accounting quality decreases for all three country clusters from the pre-adoption to the adoption period. In the post-adoption period, the metrics do not describe the development of accounting quality consistently. In *Figure 10*, the two ranges of potential outcomes of accounting quality in the post-adoption period for firms from ERC and SEC have been combined for ease of illustration. However, for all clusters accounting quality in the post-adoption period remains lower than in the pre-adoption period. This indicates that overall IFRS adoption did not lead to an increase in accounting quality and thus failed to reach one of its objectives. As the periods in this test

[20] The results of the five earnings management and the comparability metrics are tabulated in *Appendix 9*. The results on value relevance are not tabulated because the results do not allow any important inferences.

correctly capture the timing of IFRS adoption for all countries and the test thus better captures the effect of IFRS adoption, it can be viewed as confirmation of our initial inferences related to IFRS' first-order effects. The only major change compared to the results including the five countries is that in the first period there now is a significant difference in accounting quality between firms from ERC and SEC. In the following two periods, both tests do not indicate significant differences for the two clusters any more. This might suggest that the enforcement reforms in ERC might have mitigated the decrease in accounting quality following IFRS adoption compared to SEC. However, the composition of cluster ERC has not been changed compared to our primary sample. Hence, the difference to our previous findings is primarily driven through the change in the composition of SEC (which in the robustness check only include Australia, Hong Kong and South Africa). Moreover, again the enforcement reforms do not lead to an increase in accounting quality.

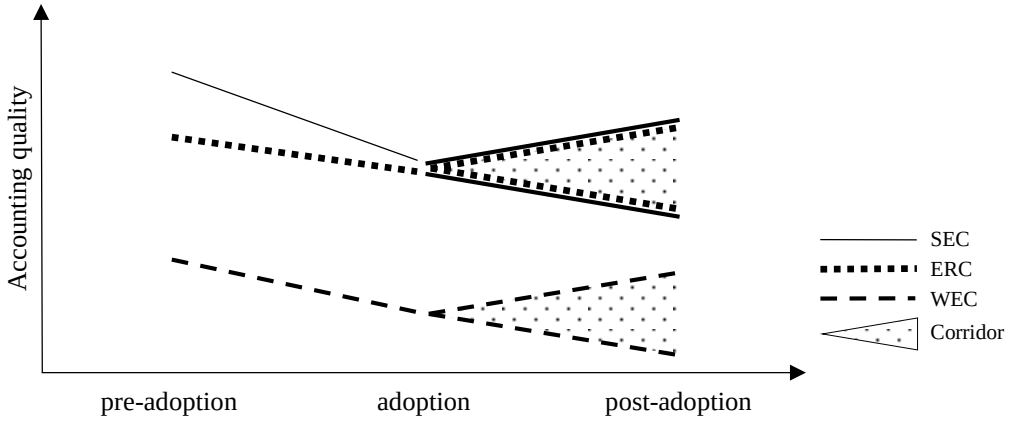

Figure 10: Accounting quality excluding IL, NZ, PK, SG, TR

7.3.3.2 Accounting comparability

Similarly, we re-calculate the results of the two accounting comparability metrics without including the five countries into the sample. *Figure 11* and *Figure 12* show the results of the robustness test for the two metrics.[21] The general picture of the results remains unchanged. As we do not test for significance we cannot evaluate whether the changes of the results compared to the primary sample are significant. However, when comparing the graphs of the robustness tests (*Figure 11* and *Figure 12*) with the results including the five countries (*Figure 4* and *Figure 5*), no big differences can be noticed. Neither the relative ranking between the clusters nor the development of accounting comparability changes. Hence, this is further evidence that the five countries that adopted IFRS at different dates than the EU countries are not driving the results and our inferences remain unchanged.

[21] However, again we do not test for significant differences between the results of each cluster pair because a bootstrapping procedure with 30 replications would have taken over two weeks.

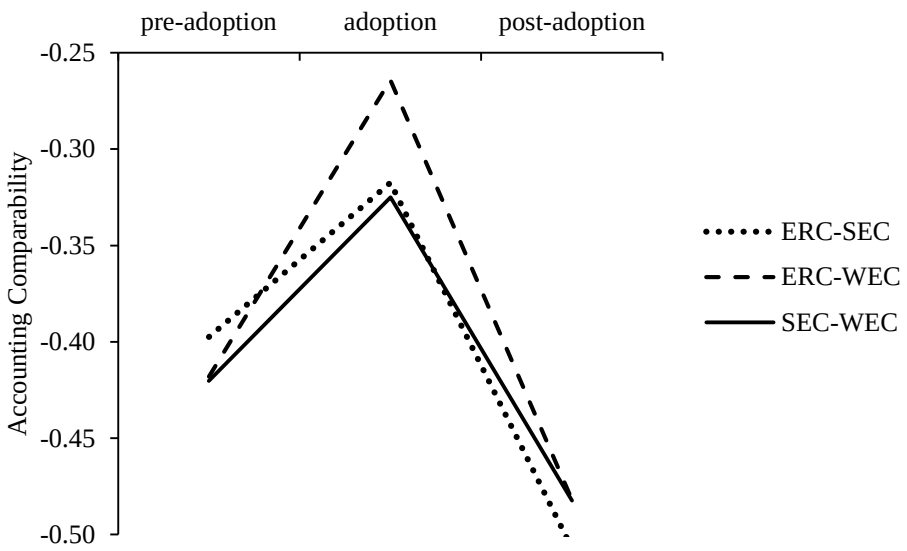

Figure 11: Return-based acc. comparability metric excluding IL, NZ, PK, SG, TR

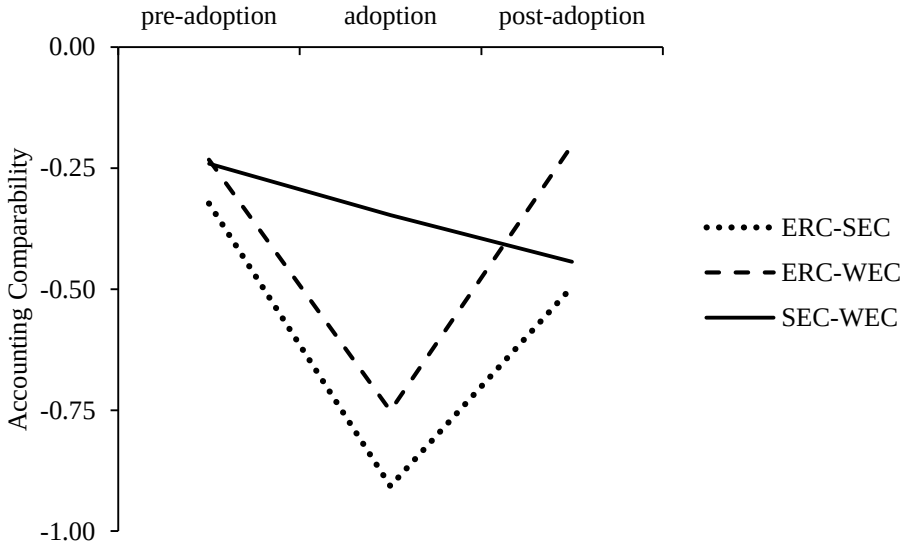

Figure 12: Cash flow-based acc. comparability metric excluding IL, NZ, PK, SG, TR

7.4 Limitations and further areas of future research

A number of caveats need to be acknowledged when interpreting our results. First, due to two reasons the composition of firms changes slightly between the periods. The first reason is that we use a non-constant sample, i.e. we do not require firm data to be fully available for the whole time period but only for sub-periods in order to include a firm into our sample. As we focus on a twelve-year time period only including firms with observations for the whole period would decrease the sample size and thereby the meaningfulness of our findings. Moreover, data for the year 2012 was still not available in DataStream for some of our sample firms. Therefore, each of the first two periods includes firms with observations for the whole respective period and in the third period this restriction is relaxed to include as well firms with observations for three of the four years.

Second, unlike Barth *et al.* (2008) and Ahmed *et al.* (2012) we do not include a benchmark control group of firms that did not adopt IFRS into our analysis. This is mainly due to the fact that these two studies analyze differences between IFRS adopters and non-adopters while we are analyzing differences between the different clusters of IFRS adopters. However, it would be interesting to see how accounting quality and comparability developed in the sample period for firms applying e.g. US GAAP. This would allow us to see whether the observed developments are restricted to our sample countries or representative of global trends. Nonetheless, adding a further group of countries to our analysis would increase the complexity of analysis and duration of calculations exponentially: For example, if you distinguish two periods and clusters, you have to compare four differences between clusters for distinct periods, four differences between periods for distinct clusters and four differences-in-differences. This is the case for Barth *et al.* (2008) and Ahmed *et al.* (2012). In our case, we analyzed three periods and three clusters, thus having two compare nine differences between clusters for distinct periods, nine differences between periods for distinct clusters and also nine differences-in-differences.

In addition, our sample includes both voluntary and mandatory adopters of IFRS. As the share of voluntary adopters in the sample is low (<5%), including them into the analysis should not constitute a problem. Still, the results might be biased to a small extent due to self-selection concerns. On the other hand, Daske *et al.* (2011) point out that it is difficult to consistently filter out voluntary adopters because the Worldscope classifications of which accounting standards are used contain a large number of inconsistencies.

Furthermore, the metrics used in our tests do not capture all relevant aspects of accounting quality and comparability (see also Brüggemann *et al.*, 2012). For example, IFRS require an increase in disclosures compared to many former local GAAP which might also lead to increases in accounting quality that are not captured by our metrics. Further research might therefore investigate other facets of accounting quality and comparability. Finally, we measure accounting quality and comparability on a highly aggregated level (see also Brüggemann *et al.*, 2012). Consequently, isolating the effects of single factors presents a challenge. This problem became evident when interpreting the results of the two accounting comparability metrics. As we did not aggregate the results of the different accounting quality metrics into one score of accounting quality per cluster and period, this issue is slightly less important for our analysis of accounting quality. Still, future research is required on a more micro level. For instance, it seems insightful to investigate which factors determine the quality and comparability of financial statements on a firm-level.

8 Conclusion

This study examines the role of IFRS adoption and enforcement for accounting quality and comparability. Prior research shows that the harmonization of accounting standards through the introduction of IFRS had positive effects on capital markets. However, the effects were restricted to countries characterized by certain institutional characteristics (e.g. the implementation of enforcement reforms concurrent to IFRS adoption or high governance

quality). Only few studies analyze how changes in accounting quality and comparability following IFRS adoption lead to the observed capital market effects. Their results are partially contradicting but indicate that enforcement systems have an effect on accounting quality and comparability by shaping reporting incentives.

In order to investigate how enforcement and its changes affect accounting quality and comparability, we cluster countries which adopted IFRS into three groups according to differences in their enforcement mechanisms. Furthermore, we distinguish between three time periods, namely before, during, and after IFRS adoption to examine whether observed effects on accounting quality and comparability are of temporary or permanent nature. Our findings indicate that there are robust differences in accounting quality and comparability between the three country clusters showing that enforcement does play a role for the quality and comparability of financial statements. This is in line with the results of prior research. On the other hand, our findings further report a decrease in accounting quality and comparability for all country clusters in the years immediately following IFRS adoption. This evidence suggests that the introduction of IFRS did not lead to an improvement in accounting quality and comparability as measured by our empirical constructs and therefore might have failed to reach one of its major objectives. Additionally, in particular for firms from countries which implemented substantial changes in their enforcement regimes concurrent to the adoption of IFRS, no increases neither in accounting quality nor in accounting comparability with other firms could be noticed. It therefore seems that the enforcement reforms undertaken parallel to IFRS adoption in some countries had only marginal effects on accounting quality and comparability, suggesting that the role of enforcement on accounting quality and comparability is lower than indicated by prior research.

Moreover, our thesis indicates that accounting quality and comparability might have started to increase again in more recent years. However, the metrics used in the tests do not show a consistent picture in this respect. Still, this finding might imply the existence of potential learning effects as preparers of financial statements needed some time to adjust to the new accounting standards. Further, our findings show that differences in accounting quality and comparability between country clusters continue to exist also in recent years. This is in line with the study from Leuz (2010) who argues that despite the harmonization of accounting standards, institutional differences between country clusters continue to exist and therefore also differences in reporting practices between different countries. Consequently, our findings suggest a divergence rather than convergence of reporting practices between countries following IFRS adoption.

Our thesis contributes to existing literature in three ways: First, unlike prior studies we follow a comprehensive approach of examining both accounting quality and comparability using a variety of different measures. Second, our thesis investigates a long-term time period and allows for a differentiation between permanent and temporary effects. Thus, we are able to reconcile partially contradicting findings of prior literature. Third, regarding the area of accounting comparability little empirical research has been done before because for a long time no sound measure of accounting comparability was available. Using the methodology proposed by De Franco *et al.* (2011), our thesis is one of few studies which adds to this line of research. We

find, however, that the measure suggested by De Franco *et al.* does not come without its disadvantages in that it might be affected by capital market developments which are unrelated to accounting comparability.

Nevertheless, our thesis comes about with some limitations. First, the metrics used in our tests do not capture all relevant aspects of accounting quality and comparability. For example, IFRS require higher levels of disclosures from firms compared to many former local GAAP which might also lead to increases in accounting quality that are not captured by our metrics. Further research might therefore investigate other facets of accounting quality and comparability. Second, we measure accounting quality and comparability on a highly aggregated level. Accordingly, isolating the effects of single factors presents a challenge which requires future research to focus on a micro level. For instance, it seems insightful to investigate which factors determine the quality and comparability of financial statements on a firm-level. Finally, no control group of IFRS non-adopting countries was included into our research design as this would have obviously exceeded the scope of our thesis. Nonetheless, it would be interesting to know how accounting quality and comparability developed for such countries as this would allow to differentiate further between general and IFRS-specific effects.

References

Ahmed, A. S., Neel, M. and Wang, E. (2012) Does Mandatory Adoption of IFRS Improve Accounting Quality? Preliminary Evidence, *Contemporary Accounting Research (forthcoming)*.

Amir, E., Harris, T. S. and Venuti, E. K. (1993) A Comparison of the Value-Relevance of U.S. versus Non-U.S. GAAP Accounting Measures Using Form 20-F Reconciliations, *Journal of Accounting Research*, 31(3), pp. 230-64.

Antoniou, A., Guney, Y. and Paudyal, K. (2008) The Determinants of Capital Structure: Capital Market-Oriented versus Bank-Oriented Institutions, *Journal of Financial and Quantitative Analysis*, 43(01), pp. 59-92.

Armstrong, C. S., Barth, M. E., Jagolinzer, A. D. and Riedl, E. J. (2010) Market Reaction to the Adoption of IFRS in Europe, *Accounting Review*, 85(1), pp. 31-61.

Bae, K.-H., Tan, H. and Welker, M. (2008) International GAAP Differences: The Impact on Foreign Analysts, *Accounting Review*, 83(3), pp. 593-628.

Ball, R. (2006) International Financial Reporting Standards (IFRS): pros and cons for investors, *Accounting & Business Research (Wolters Kluwer UK)*, 36, pp. 5-27.

Ball, R. (1995) Making Accounting More International: Why, how, and how far will it go?, *Journal of Applied Corporate Finance*, 8(3), pp. 19-29.

Ball, R., Kothari, S. P. and Robin, A. (2000) The effect of international institutional factors on properties of accounting earnings, *Journal of Accounting and Economics*, 29(1), pp. 1-51.

Ball, R., Robin, A. and Wu, J. S. (2003) Incentives versus standards: properties of accounting income in four East Asian countries, *Journal of Accounting and Economics*, 36(1–3), pp. 235-70.

Ball, R. and Shivakumar, L. (2005) Earnings quality in UK private firms: comparative loss recognition timeliness, *Journal of Accounting and Economics*, 39(1), pp. 83-128.

Ball, R. and Shivakumar, L. (2006) The Role of accruals in asymmetrically timely gain and loss recognition, *Journal of Accounting Research*, 44(2), pp. 207-42.

Barth, M. E., Beaver, W. H. and Landsman, W. R. (2001) The relevance of the value relevance literature for financial accounting standard setting: another view, *Journal of Accounting and Economics*, 31(1–3), pp. 77-104.

Barth, M. E., Landsman, W. R., Lang, M. and Williams, C. (2012) Are IFRS-based and US GAAP-based accounting amounts comparable?, *Journal of Accounting & Economics*, 54(1), p. 68.

Barth, M. E., Landsman, W. R. and Lang, M. H. (2008) International Accounting Standards and Accounting Quality, *Journal of Accounting Research*, 46(3), pp. 467-98.

Barzel, Y. (1982) Measurement Cost and the Organization of Markets, *Journal of Law and Economics*, 25(1), pp. 27-48.

Beneish, M. D., Miller, B. P. and Yohn, T. L. (2012) The Impact of Financial Reporting on Equity versus Debt Markets: Macroeconomic Evidence from Mandatory IFRS Adoption (Rochester, Rochester).

Brochet, F., Jagolinzer, A. D. and Riedl, E. J. (2012) Mandatory IFRS Adoption and Financial Statement Comparability (Rochester, Rochester).

Brown, P. and Tarca, A. (2001) Politics, Processes and the Future of Australian Accounting Standards, *Abacus*, 37(3), pp. 267-96.

Brüggemann, U., Hitz, J.-M. and Sellhorn, T. (2012) Intended and Unintended Consequences of Mandatory IFRS Adoption: A Review of Extant Evidence and Suggestions for Future Research (Rochester, Rochester).

Burgstahler, D. and Dichev, I. (1997) Earnings management to avoid earnings decreases and losses, *Journal of Accounting and Economics*, 24(1), pp. 99-126.

Burgstahler, D. C., Hail, L. and Leuz, C. (2006) The Importance of Reporting Incentives: Earnings Management in European Private and Public Firms, *Accounting Review*, 81(5), pp. 983-1016.

Byard, D., Li, Y. and Yu, Y. (2011) The Effect of Mandatory IFRS Adoption on Financial Analysts' Information Environment, *Journal of Accounting Research*, 49(1), pp. 69-96.

Carvajal, A. and Elliott, J. (2009) The Challenge of Enforcement inSecurities Markets: Mission Impossible?, *International Monetary Fund, Working Paper WP/09/168, available at* *http://www.imf.org/external/pubs/ft/wp/2009/wp09168.pdf*.

Cascino, S. and Gassen, J. (2012) Comparability Effects of Mandatory IFRS Adoption, *Working paper, forthcoming*.

Christensen, H. B., Hail, L. and Leuz, C. (2012) Mandatory IFRS Reporting and Changes in Enforcement, *Working paper, forthcoming, available under* *http://ssrn.com/abstract=2017160*.

Christensen, H. B., Lee, E. and Walker, M. (2008) Incentives or Standards: What Determines Accounting Quality Changes Around IFRS Adoption? (Rochester, Rochester).

Coase, R. H. (1960) The Problem of Social Cost, *Journal of Law and Economics*, 3, pp. 1-44.

Cormier, D., Demaria, S., Lapointe-Antunes, P. and Teller, R. (2009) First-Time Adoption of IFRS, Managerial Incentives, and Value-Relevance: Some French Evidence, *Journal of International Accounting Research*, 8(2), pp. 1-22.

Daske, H., Hail, L., Leuz, C. and Verdi, R. (2008) Mandatory IFRS Reporting around the World: Early Evidence on the Economic Consequences, *Journal of Accounting Research*, 46(5), pp. 1085-142.

Daske, H., Hail, L., Leuz, C. and Verdi, R. S. (2011) Adopting a Label: Heterogeneity in the Economic Consequences around Voluntary IAS and IFRS Adoptions (Rochester, Rochester).

De Franco, G. U. S., Kothari, S. P. and Verdi, R. S. (2011) The Benefits of Financial Statement Comparability, *Journal of Accounting Research*, 49(4), pp. 895-931.

DeFond, M., Hu, X., Hung, M. and Li, S. (2011) The impact of mandatory IFRS adoption on foreign mutual fund ownership: The role of comparability, *Journal of Accounting and Economics*, 51(3), pp. 240-58.

Demirgüç-Kunt, A. and Levine, R. (1999) Bank-Based and Market-Based Financial Systems: Cross-Country Comparisons, *World Bank Policy Working Paper No. 2143*.

Djankov, S., La Porta, R., Lopez-de-Silanes, F. and Shleifer, A. (2008) The law and economics of self-dealing, *Journal of Financial Economics*, 88(3), pp. 430-65.

Ergungor, O. E. (2004) Market- vs. bank-based financial systems: Do rights and regulations really matter?, *Journal of Banking & Finance*, 28(12), pp. 2869-87.

European Communities (2002) Regulation (EC) No 1606/2002 of the European Parliament and of the Council of 19 July 2002 on the application of international accounting standards. Official Journal of the European Communities, L 243, pp. 1-4.

Fédération des Experts Comptables Européens (2001) Enforcement mechanisms in Europe - A preliminary investigation of oversight systems *available at:* *http://www.fee.be/images/publications/accounting/Enforcement%20Mechanisms%20in%20Europe153200520952.pdf*.

Garcia Osma, B. and Pope, P. F. (2011) Strategic Balance Sheet Adjustments under First-Time IFRS Adoption and the Consequences for Earnings Quality, *Working paper, forthcoming*.

Gjesdal, F. (1981) Accounting for Stewardship, *Journal of Accounting Research*, 19(1), pp. 208-31.

Godfrey, J. M., Hodgson, A., Tarca, A., Hamilton, J. and Holmes, S. (2010) *Accounting theory* (Milton, Qld.: John Wiley).

Hail, L., Leuz, C. and Wysocki, P. (2010) Global Accounting Convergence and the Potential Adoption of IFRS by the U.S. (Part I): Conceptual Underpinnings and Economic Analysis, *Accounting Horizons*, 24(3), pp. 355-94.

Holthausen, R. W. (2003) Testing the relative power of accounting standards versus incentives and other institutional features to influence the outcome of financial reporting in an international setting, *Journal of Accounting and Economics*, 36(1–3), pp. 271-83.

International Financial Reporting Standards Foundation (2010) The Conceptual Framework for Financial Reporting.

International Financial Reporting Standards Foundation (2013) Constitution.

Jensen, M. C. and Meckling, W. H. (1976) Theory of the firm: Managerial behavior, agency costs and ownership structure, *Journal of Financial Economics*, 3(4), pp. 305-60.

Kaufmann, D., Kraay, A. and Mastruzzi, M. (2009) Governance Matters VIII: Aggregate and Individual Governance Indicators, 1996-2008 (Rochester, Rochester).

Kennedy, P. (2008) *A guide to econometrics* (Blackwell Publishing).

Knutson, P. H. and Napolitano, G. U. (1998) Criteria Employed by the AIMR Financial Accounting Policy Committee in Evaluating Financial Accounting Standards, *Accounting Horizons*, 12, pp. 170-76.

La Porta, R., Lopez-De-Silanes, F. and Shleifer, A. (2006) What Works in Securities Laws?, *Journal of Finance*, 61(1), pp. 1-32.

Land, J. and Lang, M. H. (2002) Empirical Evidence on the Evolution of International Earnings, *The Accounting Review*, 77, pp. 115-33.

Lang, M., Raedy, J. S. and Yetman, M. H. (2003) How Representative Are Firms That Are Cross-Listed in the United States? An Analysis of Accounting Quality, *Journal of Accounting Research*, 41(2), pp. 363-86.

Lang, M., Smith Raedy, J. and Wilson, W. (2006) Earnings management and cross listing: Are reconciled earnings comparable to US earnings?, *Journal of Accounting and Economics*, 42(1–2), pp. 255-83.

Lang, M. H., Maffett, M. G. and Owens, E. L. (2010) Earnings Comovement and Accounting Comparability: The Effects of Mandatory IFRS Adoption (Rochester, Rochester).

Leuz, C. (2010) Different approaches to corporate reporting regulation: how jurisdictions differ and why, *Accounting and Business Research*, 40(3), pp. 229-56.

Leuz, C., Nanda, D. and Wysocki, P. D. (2003) Earnings management and investor protection: an international comparison, *Journal of Financial Economics*, 69(3), pp. 505-27.

Leuz, C. and Verrecchia, R. E. (2000) The Economic Consequences of Increased Disclosure, *Journal of Accounting Research*, 38(3), pp. 91-124.

Li, S. (2010) Does Mandatory Adoption of International Financial Reporting Standards in the European Union Reduce the Cost of Equity Capital?, *The Accounting Review*, 85(2), pp. 607-36.

Maddala, G. S. (1991) A Perspective on the Use of Limited-Dependent and Qualitative Variables Models in Accounting Research, *The Accounting Review*, 66(4), pp. 788-807.

Nobes, C. (2006) The survival of international differences under IFRS: towards a research agenda, *Accounting & Business Research (Wolters Kluwer UK)*, 36(3), pp. 233-45.

North, D. C. (1990) *Institutions, institutional change and economic performance* (Cambridge: Cambridge University Press).

Olson, M. (1971) *The logic of collective action : public goods and the theory of groups* (Cambridge, Mass.: Harvard University Press).

Ormrod, P. and Taylor, P. (2004) The Impact of the Change to International Accounting Standards on Debt Covenants: A UK Perspective, *Accounting in Europe*, 1, pp. 71-94.

Prather-Kinsey, J. (2006) Developing countries converging with developed-country accounting standards: Evidence from South Africa and Mexico, *The International Journal of Accounting*, 41(2), pp. 141-62.

Preiato, J. P., Brown, P. R. and Tarca, A. (2013) Mandatory Adoption of IFRS and Analysts' Forecasts: How Much Does Enforcement Matter?, *Working paper, forthcoming*.

Runsten, M. (1998) 'The association between accounting information and stock prices', Stockholm School of Economics, Dissertation.

Samarasekera, N., Chang, M. and Tarca, A. (2012) IFRS and accounting quality: The impact of enforcement, *Working paper, forthcoming, available under http://papers.ssrn.com/sol3/papers.cfm?abstract_id=2183061*.

Scott, W. R. (2001) *Institutions and organizations* (Thousand Oaks, Calif.: Sage).

Soderstrom, N. S. and Sun, K. J. (2007) IFRS Adoption and Accounting Quality: A Review, *European Accounting Review*, 16(4), pp. 675-702.

Watts, R. L. and Zimmerman, J. L. (1986) *Positive accounting theory* (Englewood Cliffs, N.J.: Prentice-Hall).

Whittington, G. (2005) The adoption of International Accounting Standards in the European Union, *European Accounting Review*, 14(1), pp. 127-53.

Williamson, O. E. (1975) *Markets and hierarchies : analysis and antitrust implications : a study in the economics of internal organization* (New York: Free Press).

Wysocki, P. (2011) New institutional accounting and IFRS, *Accounting and Business Research*, 41(3), pp. 309-28.

Appendix

Appendix 1: Countries which adopted IFRS before 2009

Country	Adoption of mandatory IFRS reporting	Country	Adoption of mandatory IFRS reporting
Abu Dhabi	12/31/03	Luxembourg	12/31/05
Australia	12/31/05	Netherlands	12/31/05
Austria	12/31/05	New Zealand	12/31/07
Belgium	12/31/05	Norway	12/31/05
Czech Republic	12/31/05	Pakistan	12/31/07
Denmark	12/31/05	Philippines	12/31/05
Estonia	12/31/05	Poland	12/31/05
Finland	12/31/05	Portugal	12/31/05
France	12/31/05	Singapore	12/31/03
Germany	12/31/05	Slovakia	12/31/05
Greece	12/31/05	Slovenia	12/31/05
Hong Kong	12/31/05	South Africa	12/31/05
Hungary	12/31/05	Spain	12/31/05
Iceland	12/31/05	Sweden	12/31/05
Ireland	12/31/05	Switzerland	12/31/05
Israel	12/31/08	Turkey	12/31/06
Italy	12/31/05	United Kingdom	12/31/05
Lithuania	12/31/05		

Appendix 2: Sample countries

Australia	Germany	New Zealand	South Africa
Austria	Greece	Norway	Spain
Belgium	Hong Kong	Pakistan	Sweden
Denmark	Israel	Philippines	Switzerland
Finland	Italy	Portugal	Turkey
France	Netherlands	Singapore	United Kingdom

Appendix 3: Description of variables of institutional characteristics

Variable	Description
Security regulation	
Disclosure requirements	level of disclosure requirements in securities offerings; La Porta *et al.* (2006)
Liability standard	the arithmetic mean of the liability standards for issuers, its directors, distributors, and accountants; La Porta *et al.* (2006)
Public enforcement	summary index of several subindices on public enforcement of securities regulation (supervisor characteristics index, rule-making power index, investigative powers index, orders index, and criminal index); La Porta *et al.* (2006)
Protection of outsiders against self-dealing by insiders	
Ex-ante control of self-dealing	average of requirements for approval by disinterested shareholders and ex-ante disclosure; Djankov *et al.* (2008)
Ex-post control of self-dealing	average of disclosure in periodic filings and ease of proving wrongdoing; Djankov *et al.* (2008)
Public enforcement of anti self-dealing provision	measures available fines and sanctions to the public enforcer; Djankov *et al.* (2008)

Appendix 4: Institutional characteristics (Djankov *et al.*, 2008; La Porta *et al.*, 2006)

	Securities Regulation			Self-Dealing		
Country name	**Disclosure require-ments**	**Liability standard**	**Public enforce-ment**	**Ex-ante control**	**Ex-post control**	**Public enforce-ment**
AUSTRALIA	0.75	0.66	0.9	0.89	0.63	0.5
AUSTRIA	0.25	0.11	0.17	0	0.43	1
BELGIUM	0.42	0.44	0.15	0.39	0.7	0.5
DENMARK	0.58	0.55	0.37	0.25	0.68	0.75
FINLAND	0.5	0.66	0.32	0.14	0.77	0
FRANCE	0.75	0.22	0.77	0.08	0.68	0.5
GERMANY	0.42	0	0.22	0.14	0.43	1
GREECE	0.33	0.5	0.32	0.08	0.35	0.5
HONGKONG	0.92	0.66	0.87	1	0.93	0
ISRAEL	0.67	0.66	0.63	0.5	0.95	1
ITALY	0.67	0.67	0.48	0.17	0.68	0
NETHERLANDS	0.5	0.89	0.47	0.06	0.35	0
NEW ZEALAND	0.67	0.44	0.33	1	0.9	0
NORWAY	0.58	0.39	0.32	0.42	0.43	1
PAKISTAN	0.58	0.39	0.58	0.17	0.65	0.75
PHILIPPINES	0.83	1	0.83	0.06	0.38	0
PORTUGAL	0.42	0.66	0.58	0.14	0.75	1
SINGAPORE	1	0.66	0.87	1	1	1
SOUTHAFRICA	0.83	0.66	0.25	1	0.63	0
SPAIN	0.5	0.66	0.33	0.22	0.52	1
SWEDEN	0.58	0.28	0.5	0.17	0.5	1
SWITZERLAND	0.67	0.44	0.33	0.08	0.45	0.75
TURKEY	0.5	0.22	0.63	0.33	0.52	0
UNITED KINGDOM	0.83	0.66	0.68	1	0.9	0

Appendix 5: Metrics of accounting quality

Metric number	Accounting quality metrics	Models	Better accounting quality
	I. Earnings management		
1.1	**Earnings smoothing** Variance in the residuals of ΔNI model	$\Delta NI_{it} = \alpha_0 + \alpha_1 SIZE_{it} + \alpha_2 GROWTH_{it} + \alpha_3 EISSUE_{it} + \alpha_4 LEV_{it} + \alpha_5 DISSUE_{it} + \alpha_6 TURN_{it} + \alpha_7 CF_{it} + \alpha_8 AUD_{it} + \alpha_9 NUMEX_{it} + \alpha_{10} XLIST_{it} + \alpha_{11} CLOSE_{it} + \varepsilon_{it}$	Higher variance
1.2	Variance of residuals from the ΔNI model scaled by the variance of residuals from the ΔCF model	$\Delta CF_{it} = \alpha_0 + \alpha_1 SIZE_{it} + \alpha_2 GROWTH_{it} + \alpha_3 EISSUE_{it} + \alpha_4 LEV_{it} + \alpha_5 DISSUE_{it} + \alpha_6 TURN_{it} + \alpha_7 CF_{it} + \alpha_8 AUD_{it} + \alpha_9 NUMEX_{it} + \alpha_{10} ADR_{it} + \alpha_{11} CLOSE_{it} + \varepsilon_{it}$ $\textit{Ratio of variance of residuals of change in net income}_{it} \textit{ by variance of residuals of change in cash flow}_{it} = \left(\frac{\Delta NI*}{\Delta CF*}\right)_{it}$	Higher variance
1.3	Correlation between the residuals of the CF and ACC models	$CF_{it} = \alpha_0 + \alpha_1 SIZE_{it} + \alpha_2 GROWTH_{it} + \alpha_3 EISSUE_{it} + \alpha_4 LEV_{it} + \alpha_5 DISSUE_{it} + \alpha_6 TURN_{it} + \alpha_7 AUD_{it} + \alpha_9 ADR_{it} + \alpha_{10} CLOSE_{it} + \varepsilon_{it}$ $ACC_{it} = \alpha_0 + \alpha_1 SIZE_{it} + \alpha_2 GROWTH_{it} + \alpha_3 EISSUE_{it} + \alpha_4 LEV_{it} + \alpha_5 DISSUE_{it} + \alpha_6 TURN_{it} + \alpha_7 AUD_{it} + \alpha_8 NUMEX_{it} + \alpha_9 ADR_{it} + \alpha_{10} CLOSE_{it} + \varepsilon_{it}$	Lower negative correlation
1.4	**Managing towards earnings targets** Coefficient of $SPOS$ across clusters (CD) and periods (PD)	$CD(0,1)_{it} = \alpha_0 + \alpha_1 SPOS_{it} + \alpha_2 SIZE_{it} + \alpha_3 GROWTH_{it} + \alpha_4 EISSUE_{it} + \alpha_5 LEV_{it} + \alpha_6 DISSUE_{it} + \alpha_7 TURN_{it} + \alpha_8 CF_{it} + \alpha_9 AUD_{it} + \alpha_{10} NUMEX_{it} + \alpha_{11} ADR_{it} + \alpha_{12} CLOSE_{it} + \varepsilon_{it}$ $PD(0,1)_{it} = \alpha_0 + \alpha_1 SPOS_{it} + \alpha_2 SIZE_{it} + \alpha_3 GROWTH_{it} + \alpha_4 EISSUE_{it} + \alpha_5 LEV_{it} + \alpha_6 DISSUE_{it} + \alpha_7 TURN_{it} + \alpha_8 CF_{it} + \alpha_9 AUD_{it} + \alpha_{10} NUMEX_{it} + \alpha_{11} ADR_{it} + \alpha_{12} CLOSE_{it} + \varepsilon_{it}$	Sign of the coeffient of $SPOS$
1.5	**Timeliness of loss recognition** Coefficient of $LNNI$ across clusters (CD) and periods (PD)	$CD(0,1)_{it} = \alpha_0 + \alpha_1 LNNI_{it} + \alpha_2 SIZE_{it} + \alpha_3 GROWTH_{it} + \alpha_4 EISSUE_{it} + \alpha_5 LEV_{it} + \alpha_6 DISSUE_{it} + \alpha_7 TURN_{it} + \alpha_8 CF_{it} + \alpha_9 AUD_{it} + \alpha_{10} NUMEX_{it} + \alpha_{11} ADR_{it} + \alpha_{12} CLOSE_{it} + \varepsilon_{it}$ $PD(0,1)_{it} = \alpha_0 + \alpha_1 LNNI_{it} + \alpha_2 SIZE_{it} + \alpha_3 GROWTH_{it} + \alpha_4 EISSUE_{it} + \alpha_5 LEV_{it} + \alpha_6 DISSUE_{it} + \alpha_7 TURN_{it} + \alpha_8 CF_{it} + \alpha_9 AUD_{it} + \alpha_{10} NUMEX_{it} + \alpha_{11} ADR_{it} + \alpha_{12} CLOSE_{it} + \varepsilon_{it}$	Sign of the coeffient of $LNNI$

	II. Value relevance		
2.1	*Adjusted* R^2 of regression of P^*	$P^*_{it} = \beta_0 + \beta_1 BVEPS_{it} + \beta_2 NIPS_{it} + \varepsilon_{it}$	Higher *adjusted* R^2
2.2	*Adjusted* R^2 of regression of $(NI/P)^*$ for 'positive' subsample	$\left(\frac{NI}{P}\right)^*_{it} = \beta_0 + \beta_1 RETURN_{it} + \varepsilon_{it}$	Higher *adjusted* R^2
2.3	*Adjusted* R^2 of regression of $(NI/P)^*$ for 'negative' subsample	$\left(\frac{NI}{P}\right)^*_{it} = \beta_0 + \beta_1 RETURN_{it} + \varepsilon_{it}$	Higher *adjusted* R^2

ΔNI is the change in net income available to common shareholders at fiscal year-end scaled by end of fiscal year total assets. *SIZE* is the natural logarithm of market value of equity by end of fiscal year. *GROWTH* is annual percentage change in sales. *EISSUE* is annual percentage change in book value of equity. *LEV* is total liabilities divided by book value of equity at fiscal year-end. *DISSUE* is annual percentage change in total liabilities. *TURN* is annual sales divided by end of fiscal year total assets. *CF* is annual net cash flow from operating activities divided by fiscal year total assets. *AUD* is an indicator variable set to one if the firm's auditor is PricewaterhouseCoopers, KPMG, Arthur Andersen, Ernst & Young or Deloitte Touche Tohmatsu, and zero otherwise. *NUMEX* is the number of exchanges on which a firm's stock is listed. *ADR* is an indicator variable that equals one if the firm trades American depository shares. *CLOSE* is the percentage of closely held shares of the firm as reported by Worldscope. ΔCF is change in annual net cash flow from operating activities scaled by end of fiscal year total assets. *ACC* is annual net income available to common shareholders at fiscal year-end less annual cash flow from operating activities, scaled by end-of-year total assets. *CD* is an indicator variable equaling one or zero depending on the cluster membership of an observation. *PD* is an indicator variable equaling one or zero depending on the time period from which an observation stems. *SPOS* is an indicator variable equal to one for observations for which annual net income scaled by total assets is between 0 and 0.01 and set to zero otherwise. *LNNI* an indicator variable set to one for observations for which annual net income scaled by total assets is less than -0.2 and zero otherwise. P^* is the residuals from a regression where stock price six months after fiscal year end (P) is regressed on industry fixed effects. *BVEPS* is book value of equity per ordinary share at fiscal year-end. *NIPS* is net income available to common shareholders at fiscal year-end per ordinary share. $(NI/P)^*$ is the residuals from a regression where (NI/P) is regressed on country and fixed effects. *RETURN* is measured as the natural logarithm of the ratio of the stock price three months after fiscal year-end to the stock price nine months before fiscal year-end (adjusted for dividends and stock splits).

Appendix 6: Accounting comparability metric based on De Franco *et al.* (2011)

Step of calculation	Equation	Better accounting comparability
1. Estimation of the accounting function of firm *i*	$Earnings\_MVE_{it} = \alpha_i + \beta_i Return_{it} + \varepsilon_{it}$	
2. Construction of a measure of comparability of accounting information between firm *i* and *j*	$ACOMP_{ij} = -\frac{1}{16} \cdot \sum_{t-15}^{t} \left\lvert E(Earnings\_MVE)_{iit} - E(Earnings\_MVE)_{ijt} \right\rvert$	Greater values
3. Calculating the aggregated accounting comparability measures over each period, cluster pair and industry	$ACOMP_{p,ci,cj,k} = -\frac{\sum_{i,j} \left\lvert E(Earnings\_MVE)_{iip} - E(EarningsMVE)_{ijp} \right\rvert}{n_{p,ci,cj,k}}$	Greater values
4. Averaging over industries to get the final measure of accounting comparability between each cluster pair for each period	$AGG.ACOMP_{p,i,j} = \frac{\sum_{i,j} ACOMP_{p,ci,cj,k}}{n_k}$	Greater values

This appendix presents the most important steps in the calculation of the return-based accounting comparability metric based on De Franco *et al.* (2011) (metric 3.1). The cash-flow based metric (metric 3.2) is calculated analogously, differences are pointed out below. In the first step, the accounting function of firm *i* (the "mapping" of economic events, i.e. returns, into financial statements, i.e. earnings) is estimated by regressing earnings on returns. *Earnings_MVE* are net income before extraordinary items, scaled by market value of equity at the beginning of each year and *Return* is the stock price return. For metric 3.2, in this step earnings (net income before extraordinary items, scaled by lagged total assets) are regressed on cash flows (net cash flows from operating activities, scaled by lagged total assets).

In the second step, a measure of accounting comparability, $ACOMP_{ij}$, of two firms *i* and *j* is calculated as the negative value of the average absolute difference between the predicted earnings using firm *i*'s and *j*'s accounting functions. For this, earnings are predicted (*E(Earnings_MVE)*) for both

firms by holding the returns (as proxy for economic events) constant and then calculating the average difference between the two earnings predictions.

In the third step, these measures of comparability between firms are averaged for each cluster, peer-cluster, industry and period, resulting in aggregated accounting comparability values, $ACOMP_{p,ci,cj,k,}$ for each cluster pair, industry and period combination, where *p* indicates the period, *ci* the country cluster of firm *i*, *cj* the country cluster of firm *j*, *k* the industry group and $n_{p,ci,cj,k}$ the number of firm pairs with firm *i* from cluster *ci* and firm *j* from cluster *cj*.

In the last step, these measures are further aggregated by averaging over all industries, resulting in our final return-based measure of average accounting comparability, $AGG.ACOMP_{p,i,j.}$ between all three cluster pairs for each period, where n_k indicates the number of industries.

Appendix 7: Descriptive statistics for variables used for the accounting quality metrics

	pre-adoption						adoption						post-adoption					
	ERC		SEC		WEC		ERC		SEC		WEC		ERC		SEC		WEC	
	N	mean	N	mean	N	mean	N	mean	N	mean	N	mean	N	mean	N	mean	N	mean
Test variables																		
ΔNI	8,971	0.0117	9,156	0.0092	8,989	0.0066	9,622	-0.0042	11,179	-0.0134	11,029	-0.0022	6,270	0.0080	8,869	0.0059	8,498	0.0057
ΔCF	8,644	0.0131	9,065	0.0076	6,791	0.0126	9,274	0.0058	11,131	0.0041	10,018	0.0034	6,060	0.0026	8,773	-0.0020	8,238	0.0042
CF	9,493	0.0098	9,987	-0.0011	7,638	0.0414	9,830	0.0035	11,900	-0.0147	11,212	0.0316	6,228	0.0248	9,169	-0.0177	8,356	0.0362
ACC	9,477	-0.0782	9,973	-0.0604	7,624	-0.0637	9,817	-0.0510	11,886	-0.0579	11,205	-0.0318	6,225	-0.0608	9,167	-0.0528	8,346	-0.0512
SPOS	9,907	0.0540	10,297	0.0373	9,763	0.0817	10,203	0.0437	12,100	0.0349	11,902	0.0679	6,533	0.0507	9,350	0.0348	8,979	0.0691
LNNI	9,907	0.1809	10,297	0.1856	9,763	0.0912	10,203	0.1663	12,100	0.1985	11,902	0.0802	6,533	0.1358	9,350	0.1871	8,979	0.0864
RETURN	7,641	-0.0749	8,030	0.0265	8,039	-0.0313	8,143	-0.2188	10,081	-0.1280	9,628	-0.1565	4,622	0.1223	7,249	-0.0201	6,014	0.1140
NI/P	6,350	-0.0563	7,351	-0.1252	4,670	-0.0127	7,221	0.1934	7,957	-0.0340	6,070	0.0201	4,773	-0.0449	5,673	-0.0319	4,946	-0.0478
P	8,233	-3.4920	9,029	4.4380	8,413	-55.3499	8,940	0.1538	10,975	-1.5825	10,462	84.881	4,647	-3.2914	7,608	-2.9688	6,137	-41.7113
BVEPS	8,097	55.8347	8,651	2.3984	8,468	-705.89	8,803	339.018	10,775	2.2073	9,974	-482.31	6,109	165.378	8,756	2.6520	7,997	50.8279
NIPS	8,096	7.9498	8,646	0.2859	8,465	117.797	8,797	-55.905	10,771	0.3140	9,970	14.7607	6,139	325.950	8,753	-0.0038	8,065	3.7469
Control Variables																		
LEV	9,840	1.4580	10,167	0.9326	9,668	1.6771	10,159	1.3239	12,052	0.8845	11,828	1.5913	6,412	1.2545	9,336	0.7913	8,617	1.4948
GROWTH	8,666	0.1253	8,183	0.1520	8,971	0.1075	8,897	0.2024	9,231	0.1967	10,852	0.1528	5,928	0.0945	7,095	0.1178	8,652	0.0685
EISSUE	8,786	0.0650	9,284	0.1056	8,861	0.0456	9,265	0.0838	11,308	0.1258	10,567	0.0536	6,215	0.0869	8,907	0.1377	8,397	0.0501
DISSUE	8,989	0.1374	9,215	0.2397	9,010	0.1250	9,615	0.3094	11,274	0.3361	11,020	0.2081	6,275	0.1178	8,899	0.2511	8,507	0.0980
TURN	9,824	1.1478	9,996	0.8978	9,648	1.0344	10,156	0.9874	11,932	0.7811	11,808	0.9691	6,412	0.9678	9,274	0.6927	8,605	0.9220
SIZE	7,658	11.0641	8,247	11.7360	7,969	11.6385	8,206	11.1266	10,388	11.7180	9,446	12.0242	4,678	11.4068	7,538	11.6913	5,840	12.2399
CF	9,493	0.0098	9,987	-0.0011	7,638	0.0414	9,830	0.0035	11,900	-0.0147	11,212	0.0316	6,228	0.0248	9,169	-0.0177	8,356	0.0362
AUD	9,907	0.5839	10,297	0.5843	9,763	0.5885	10,203	0.5455	12,100	0.5396	11,902	0.6133	6,533	0.6009	9,350	0.5372	8,979	0.6456
ADR	9,907	0.0665	10,297	0.0794	9,763	0.0437	10,203	0.0560	12,100	0.0650	11,902	0.0333	6,533	0.0651	9,350	0.0609	8,979	0.0361
CLOSE	8,081	41.1965	8,624	49.9886	5,816	54.6971	8,167	40.1498	8,594	47.7018	6,137	53.2499	4,893	95.4765	6,077	1174.39	4,674	52.4420

This table presents descriptive statistics (number of observations and mean values) for important test and control variables used in our thesis. *ΔNI* is the change in annual net income where net income is scaled by fiscal year-end total assets; *ΔCF* is the change in annual net cash flow where net cash flow is scaled by fiscal year-end total assets; *CF* is annual net cash flow from operating activities, scaled by fiscal year-end total assets; *ACC* is annual net income at fiscal year-end less annual cash flow from operating activities, scaled by fiscal year-end total assets; *SPOS* is an indicator that equals one for observations with annual net income scaled by total assets between 0.00 and 0.01, and zero otherwise; *LNNI* is an indicator variable that equals one for observations with annual net income scaled by total assets less than -0.20, and zero otherwise; *RETURN* is annual stock return from nine months prior to three months after the firm's fiscal year-end; *NI/P* is annual net income per share, scaled by beginning-of-year stock price; *P* is stock price as of six months after fiscal year-end; *BVEPS* is book value of equity per share; *NIPS* is annual net income per share. *LEV* is fiscal year-end total liabilities divided by fiscal year-end equity book value; *GROWTH* is annual percentage change in sales; *EISSUE* is annual percentage change in common stock; *DISSUE* is annual percentage change in total liabilities; *TURN* is annual sales divided by fiscal year-end total assets, *SIZE* is the natural logarithm of fiscal year-end market value of equity in thousands of US dollars; *AUD* is an indicator variable that equals one if the firm's auditor is PricewaterhouseCoopers, KPMG, Arthur Andersen, Ernst & Young or Deloitte Touche Tohmatsu, and zero otherwise; *ADR* is an indicator variable that equals one if the firm trades American depository shares, and zero otherwise; *CLOSE* is the percentage of closely held shares of the firm as reported by Worldscope.

Appendix 8: Robustness test excluding the UK

			Cluster(s)						
			Absolute values			Differences			
			ERC	SEC	WEC	ERC-SEC	ERC-WEC	SEC-WEC	
Period(s)	Absolute values	1	0.0145	0.0182	0.0093	-0.0037***	0.0052***	0.0089***	
		2	0.0068	0.0166	0.0061	-0.0098***	0.0007***	0.0105***	
		3	0.0086	0.0165	0.0074	-0.0078***	0.0012*	0.0091***	
	Differences	(1)-(2)	0.0076***	0.0016*	0.0032***	0.006***	0.0044***	-0.0016***	Difference-in-differences
		(1)-(3)	0.0058***	0.0017***	0.0019***	0.0041***	0.0039***	-0.0002***	
		(2)-(3)	-0.0018**	0.0001*	-0.0013***	-0.0019***	-0.0005	0.0014***	
						Differences-in-differences			

This table presents part of the results of the robustness test where the UK has been left out of the sample. It shows the results of the first earnings management metric - the variance of the residuals from a regression of the change in net income on the control variables from equation 1 - for each period and cluster (upper-left segment), the differences between clusters for distinct periods (upper-right segment), the differences between periods for distinct clusters (lower-left segment) and the differences-in-differences between periods and clusters (lower-right segment). The differences-in-differences compare the differences between periods over clusters. We number the periods for ease of presentation where period 1 is the pre-adoption period, period 2 the adoption period and finally, period 3 the post-adoption period. Values are rounded to four decimal places. We test significance using a percentile-bootstrapping procedure. Thus, we get the empirical distribution of the differences of the metric. *, ** and *** indicate significance at the p<10%, p<5% and p<1% level, respectively.

Metric 1.1: Variance of ΔNI* (Panel 1)

			Cluster(s)						
			Absolute values			Differences			
			ERC	SEC	WEC	ERC-SEC	ERC-WEC	SEC-WEC	
Period(s)	Absolute values	1	1.7175	1.9188	1.5471	-0.2013	0.17034**	0.3717***	
		2	1.3417	1.8893	1.4616	-0.5476***	-0.12	0.4277***	
		3	1.3892	1.8860	1.2680	-0.4968***	0.1212	0.6180***	
	Differences	(1)-(2)	0.3758***	0.0295	0.0855	0.3463***	0.2903***	-0.056***	Difference-in-differences
		(1)-(3)	0.3283***	0.0328**	0.2791***	0.2954***	0.0491***	-0.2463***	
		(2)-(3)	-0.0475	0.0033	0.1937	-0.0509***	-0.2412***	-0.1903***	
						Differences-in-differences			

This table presents part of the results of the robustness test where the UK has been left out of the sample. It shows the results of the second earnings management metric - the ratio from equation 3 between the variance of the residuals from a regression of the change in net income on the control variables and the variance of the residuals from a regression of the change in cash flow on the control variables - for each period and cluster (upper-left segment), the differences between clusters for distinct periods (upper-right segment), the differences between periods for distinct clusters (lower-left segment) and the differences-in-differences between periods and clusters (lower-right segment). The differences-in-differences compare the differences between periods over clusters. We number the periods for ease of presentation where period 1 is the pre-adoption period, period 2 the adoption period and finally, period 3 the post-adoption period. Values are rounded to four decimal places. We test significance using a percentile-bootstrapping procedure. Thus, we get the empirical distribution of the differences of the metric. *, ** and *** indicate significance at the p<10%, p<5% and p<1% level, respectively.

Metric 1.2: Variance of ΔNI*/ΔCF* (Panel 2)

			Cluster(s)						
			Absolute values			Differences			
			ERC	SEC	WEC	ERC-SEC	ERC-WEC	SEC-WEC	
Period(s)	Absolute values	1	-0.2817	-0.0737	-0.2972	-0.2080***	0.0155	0.2235***	
		2	-0.3539	-0.1653	-0.4066	-0.1887***	0.0526	0.2413***	
		3	-0.3275	-0.2576	-0.4256	-0.0699*	0.098	0.1680***	
	Differences	(1)-(2)	0.0723	0.0916***	0.1094***	-0.0193	-0.0371***	-0.0178***	Difference-in-differences
		(1)-(3)	0.0458*	0.1839***	0.1284***	-0.1381***	-0.0826***	0.0555***	
		(2)-(3)	-0.0265	0.0923***	0.019	-0.1188***	-0.0455	0.0733***	
						Differences-in-differences			

This table presents part of the results of the robustness test where the UK has been left out of the sample. It shows the results of the third earnings management metric - the correlation between the residuals from a regression of cash flows on the control variables from equation 4 and the residuals from a regression of accruals on the control variables from equation 5 - for each period and cluster (upper-left segment), the differences between clusters for distinct periods (upper-right segment), the differences between periods for distinct clusters (lower-left segment) and the differences-in-differences between periods and clusters (lower-right segment). The differences-in-differences compare the difference between periods over clusters. We number the periods for ease of presentation where period 1 is the pre-adoption period, period 2 the adoption period and finally, period 3 the post-adoption period. Values are rounded to four decimal places. We test significance using a percentile-bootstrapping procedure. Thus, we get the empirical distribution of the differences of the metric. *, ** and *** indicate significance at the p<10%, p<5% and p<1% level, respectively.

Metric 1.3: Correlation between accruals and cash flows (Panel 3)

			Cluster(s)					
						Coefficients		
			ERC	SEC	WEC	ERC-SEC	ERC-WEC	SEC-WEC
Period(s)		1				-0.2661*	0.2340*	0.4279***
		2				-0.1277	0.0687	0.3076***
		3				-0.3199**	0.0159	0.4616***
	Coefficients	(1)-(2)	-0.1283	-0.1493	-0.2915***			
		(1)-(3)	0.0731	-0.1646*	-0.0611			
		(2)-(3)	0.2539	0.0016	0.1996*			

This table presents part of the results of the robustness test where the UK has been left out of the sample. It shows the results of the fourth earnings management metric - the coefficient on small positive earnings (SPOS) in the regressions given in equations 6 and 7. In the upper-right segment there are the results of the metric between clusters for distinct periods, in the lower-left segment the results between periods for distinct clusters. Negative values imply that in the cluster comparisons (e.g. ERC-SEC) firms from the first-mentioned cluster (i.e. here ERC firms) are more likely to report small positive earnings and that in the period comparisons (e.g. 1-2) firms are more likely to report small positive earnings in the first-mentioned period (i.e. here in the first period). We number the periods for ease of presentation where period 1 is the pre-adoption period, period 2 the adoption period and finally, period 3 the post-adoption period. Values are rounded to four decimal places. We test significance using a Wald-test, the values for the Wald-statistics are not tabulated. *, ** and *** indicate significance at the p<10%, p<5% and p<1% level, respectively.

Metric 1.4: Small positive earnings (Panel 4)

			Cluster(s)					
						Coefficients		
			ERC	SEC	WEC	ERC-SEC	ERC-WEC	SEC-WEC
Period(s)		1				0.1893	-0.3878***	-0.4573***
		2				1.1221***	0.0298	-1.0880***
		3				0.3514	-0.4448**	-0.6306***
	Coefficients	(1)-(2)	-1.0254***	-0.1545*	-0.7523***			
		(1)-(3)	-0.4844***	-0.3552***	-0.5899***			
		(2)-(3)	0.5498**	-0.1843***	0.1069			

This table presents part of the results of the robustness test where the UK has been left out of the sample. It shows the results of the fifth earnings management metric - the coefficient on large negative net income (LNNI) in the regressions given in equations 8 and 9. In the upper-right segment there are the results of the metric between clusters for distinct periods, in the lower-left segment the results between periods for distinct clusters. Negative values imply that in the cluster comparisons (e.g. ERC-WEC) firms from the first-mentioned cluster (i.e. here ERC firms) are more likely to recognize large losses and that in the period comparisons (e.g. 1-2) firms are more likely to recognize large losses in the first-mentioned period (i.e. here in the first period). We number the periods for ease of presentation where period 1 is the pre-adoption period, period 2 the adoption period and finally, period 3 the post-adoption period. Values are rounded to four decimal places. We test significance using a Wald-test, the values for the Wald-statistics are not tabulated. *, ** and *** indicate significance at the p<10%, p<5% and p<1% level, respectively.

Metric 1.5: Large Negative Net Income (Panel 5)

		Cluster pair		
		ERC vs. SEC	ERC vs. WEC	SEC vs. WEC
Period(s)	1	-0.4084	-0.4469	-0.4406
	2	-0.3221	-0.2761	-0.3237
	3	-0.4787	-0.4687	-0.4796

This table presents part of the results of the robustness test where the UK has been left out of the sample. It shows the results of the first, return-based accounting comparability metric from equation 23 for the aggregated average accounting comparability between each cluster pair for each period. We number the periods for ease of presentation where period 1 is the pre-adoption period, period 2 the adoption period and finally, period 3 the post-adoption period. Values are rounded to four decimal places. No tests for significant differences have been made due to the long duration of the percentile-bootstrapping calculations.

Metric 3.1: *AGG.ACOMP* (Panel 7)

		Cluster pair		
		ERC vs. SEC	ERC vs. WEC	SEC vs. WEC
Period(s)	1	-0.2494	-0.1748	-0.2278
	2	-0.8784	-0.6688	-0.3224
	3	-0.4204	-0.1780	-0.4053

This table presents part of the results of the robustness test where the UK has been left out of the sample. It shows the results of the second, cash flow-based accounting comparability metric from equation 24 for the aggregated average accounting comparability between each cluster pair for each period. We number the periods for ease of presentation where period 1 is the pre-adoption period, period 2 the adoption period and finally, period 3 the post-adoption period. Values are rounded to four decimal places. No tests for significant differences have been made due to the long duration of the percentile-bootstrapping calculations.

Metric 3.2: *AGG.CFCOMP* (Panel 7)

Appendix 9: Robustness test excluding HK, SE, TR

			Cluster(s)						
			Absolute values			Differences			
			ERC	SEC	WEC	ERC-SEC	ERC-WEC	SEC-WEC	
Period(s)	Absolute values	1	0.0196	0.0201	0.0089	-0.0005	0.0107***	0.0111***	
		2	0.0129	0.0179	0.0059	-0.0050***	0.0070***	0.0120***	
		3	0.0149	0.0182	0.0072	-0.0033***	0.0076***	0.0109***	
	Differences	(1)-(2)	0.0067***	0.0022**	0.0030***	0.0045***	0.0037***	-0.0008***	Difference-in-differences
		(1)-(3)	0.0047***	0.0019**	0.0017**	0.0028***	0.003***	0.0002***	
		(2)-(3)	-0.0020**	-0.0003	-0.0013***	-0.0017***	-0.0007	0.001***	
									Differences-in-differences

This table presents part of the results of the robustness test where those countries have been left out of the sample that implemented substantive changes in their enforcement systems subsequent to IFRS adoption (Sweden, Hong Kong and Turkey). It shows the results of the first earnings management metric - the variance of the residuals from a regression of the change in net income on the control variables from equation 1 - for each period and cluster (upper-left segment), the differences between clusters for distinct periods (upper-right segment), the differences between periods for distinct clusters (lower-left segment) and the differences-in-differences between periods and clusters (lower-right segment). The differences-in-differences compare the differences between periods over clusters. We number the periods for ease of presentation where period 1 is the pre-adoption period, period 2 the adoption period and finally, period 3 the post-adoption period. Values are rounded to four decimal places. We test significance using a percentile-bootstrapping procedure. Thus, we get the empirical distribution of the differences of the metric. *, ** and *** indicate significance at the p<10%, p<5% and p<1% level, respectively.

Metric 1.1: Variance of ΔNI* (Panel 1)

			Cluster(s)						
			Absolute values			Differences			
			ERC	SEC	WEC	ERC-SEC	ERC-WEC	SEC-WEC	
Period(s)	Absolute values	1	1.9997	1.9934	1.5695	0.0063	0.4302***	0.4239***	
		2	1.6504	1.8948	1.4251	-0.2445***	0.2253**	0.4697***	
		3	1.8211	1.9482	1.3374	-0.1271	0.4837***	0.6107***	
	Differences	(1)-(2)	0.3493***	0.0985	0.1444	0.2508***	0.2049***	-0.0458***	Difference-in-differences
		(1)-(3)	0.1786***	0.045	0.2320*	0.1334***	-0.0535***	-0.1868***	
		(2)-(3)	-0.1707	-0.0533	0.0877	-0.1174***	-0.2584***	-0.141***	
									Differences-in-differences

This table presents part of the results of the robustness test where those countries have been left out of the sample that implemented substantive changes in their enforcement systems subsequent to IFRS adoption (Sweden, Hong Kong and Turkey). It shows the results of the second earnings management metric - the ratio from equation 3 between the variance of the residuals from a regression of the change in net income on the control variables and the variance of the residuals from a regression of the change in cash flow on the control variables - for each period and cluster (upper-left segment), the differences between clusters for distinct periods (upper-right segment), the differences between periods for distinct clusters (lower-left segment) and the differences-in-differences between periods and clusters (lower-right segment). The differences-in-differences compare the differences between periods over clusters. We number the periods for ease of presentation where period 1 is the pre-adoption period, period 2 the adoption period and finally, period 3 the post-adoption period. Values are rounded to four decimal places. We test significance using a percentile-bootstrapping procedure. Thus, we get the empirical distribution of the differences of the metric. *, ** and *** indicate significance at the p<10%, p<5% and p<1% level, respectively.

Metric 1.2: Variance of ΔNI*/ΔCF* (Panel 2)

			Cluster(s)						
			Absolute values			Differences			
			ERC	SEC	WEC	ERC-SEC	ERC-WEC	SEC-WEC	
Period(s)	Absolute values	1	-0.1242	-0.0446	-0.2944	-0.0796**	0.1701***	0.2498***	
		2	-0.1491	-0.1113	-0.4058	-0.0378	0.2567***	0.2945***	
		3	-0.1366	-0.1784	-0.4242	0.0419	0.2877***	0.2458***	
	Differences	(1)-(2)	0.0249	0.0667*	0.1115**	-0.0418***	-0.0866***	-0.0448***	Difference-in-differences
		(1)-(3)	0.0123	0.1339***	0.1299***	-0.1215***	-0.1176***	0.004***	
		(2)-(3)	-0.0126	0.0672**	0.018	-0.0797***	-0.031***	0.0487***	
									Differences-in-differences

This table presents part of the results of the robustness test where those countries have been left out of the sample that implemented substantive changes in their enforcement systems subsequent to IFRS adoption (Sweden, Hong Kong and Turkey). It shows the results of the third earnings management metric - the correlation between the residuals from a regression of cash flows on the control variables from equation 4 and the residuals from a regression of accruals on the control variables from equation 5 - for each period and cluster (upper-left segment), the differences between clusters for distinct periods (upper-right segment), the differences between periods for distinct clusters (lower-left segment) and the differences-in-differences between periods and clusters (lower-right segment). The differences-in-differences compare the difference between periods over clusters. We number the periods for ease of presentation where period 1 is the pre-adoption period, period 2 the adoption period and finally, period 3 the post-adoption period. Values are rounded to four decimal places. We test significance using a percentile-bootstrapping procedure. Thus, we get the empirical distribution of the differences of the metric. *, ** and *** indicate significance at the p<10%, p<5% and p<1% level, respectively.

Metric 1.3: Correlation between accruals and cash flows (Panel 3)

			Cluster(s)					
						Coefficients		
			ERC	SEC	WEC	ERC-SEC	ERC-WEC	SEC-WEC
Period(s)		1				-0.0968	0.3358***	0.4833***
		2				-0.0974	0.3881***	0.6892***
		3				-0.4315***	0.2405**	0.9215***
	Coefficients	(1)-(2)	-0.2177**	-0.2877**	-0.2624**			
		(1)-(3)	0.0469	-0.3035**	-0.0286			
		(2)-(3)	0.2833**	-0.0103	0.1956*			

This table presents part of the results of the robustness test where those countries have been left out of the sample that implemented substantive changes in their enforcement systems subsequent to IFRS adoption (Sweden, Hong Kong and Turkey). It shows the results of the fourth earnings management metric - the coefficient on small positive earnings (SPOS) in the regressions given in equations 6 and 7. In the upper-right segment there are the results of the metric between clusters for distinct periods, in the lower-left segment the results between periods for distinct clusters. Negative values imply that in the cluster comparisons (e.g. ERC-SEC) firms from the first-mentioned cluster (i.e. here ERC firms) are more likely to report small positive earnings and that in the period comparisons (e.g. 1-2) firms are more likely to report small positive earnings in the first-mentioned period (i.e. here in the first period). We number the periods for ease of presentation where period 1 is the pre-adoption period, period 2 the adoption period and finally, period 3 the post-adoption period. Values are rounded to four decimal places. We test significance using a Wald-test, the values for the Wald-statistics are not tabulated. *, ** and *** indicate significance at the p<10%, p<5% and p<1% level, respectively.

Metric 1.4: Small positive earnings (Panel 4)

			Cluster(s)					
						Coefficients		
			ERC	SEC	WEC	ERC-SEC	ERC-WEC	SEC-WEC
Period(s)		1				0.1536	-0.5704***	-0.5567***
		2				0.4362***	-0.5524***	-1.1326***
		3				0.3269***	-0.4427***	-0.6718***
	Coefficients	(1)-(2)	-0.535***	-0.1626*	-0.6269***			
		(1)-(3)	-0.5427***	-0.3436***	-0.4756***			
		(2)-(3)	0.0125	-0.2097*	0.1004			

This table presents part of the results of the robustness test where those countries have been left out of the sample that implemented substantive changes in their enforcement systems subsequent to IFRS adoption (Sweden, Hong Kong and Turkey). It shows the results of the fifth earnings management metric - the coefficient on large negative net income (LNNI) in the regressions given in equations 8 and 9. In the upper-right segment there are the results of the metric between clusters for distinct periods, in the lower-left segment the results between periods for distinct clusters. Negative values imply that in the cluster comparisons (e.g. ERC-WEC) firms from the first-mentioned cluster (i.e. here ERC firms) are more likely to recognize large losses and that in the period comparisons (e.g. 1-2) firms are more likely to recognize large losses in the first-mentioned period (i.e. here in the first period). We number the periods for ease of presentation where period 1 is the pre-adoption period, period 2 the adoption period and finally, period 3 the post-adoption period. Values are rounded to four decimal places. We test significance using a Wald-test, the values for the Wald-statistics are not tabulated. *, ** and *** indicate significance at the p<10%, p<5% and p<1% level, respectively.

Metric 1.5: Large Negative Net Income (Panel 5)

Appendix 10: Robustness test excluding IL, NZ, PK, SG, TR

			Cluster(s)						
			Absolute values			Differences			
			ERC	SEC	WEC	ERC-SEC	ERC-WEC	SEC-WEC	
Period(s)	Absolute values	1	0.0212	0.0247	0.0104	-0.0035***	0.0108***	0.0143***	
		2	0.0138	0.0215	0.0062	-0.0077***	0.0076***	0.0153***	
		3	0.0160	0.0215	0.0079	-0.0055***	0.0082***	0.0136***	
	Differences	(1)-(2)	0.0074***	0.0032***	0.0043***	0.0042***	0.0031***	-0.001***	Difference-in-differences
		(1)-(3)	0.0052***	0.0032***	0.0026***	0.002***	0.0026***	0.0006***	
		(2)-(3)	-0.0022*	0.00	-0.0017***	-0.0022***	-0.0005***	0.0017***	
							Differences-in-differences		

This table presents part of the results of the robustness test where countries with IFRS adoption dates differing from the end of 2005 have been left out of the sample (Israel, New Zealand, Pakistan, Singapore and Turkey). It shows the results of the first earnings management metric - the variance of the residuals from a regression of the change in net income on the control variables from equation 1 - for each period and cluster (upper-left segment), the differences between clusters for distinct periods (upper-right segment), the differences between periods for distinct clusters (lower-left segment) and the differences-in-differences between periods and clusters (lower-right segment). The differences-in-differences compare the differences between periods over clusters. We number the periods for ease of presentation where period 1 is the pre-adoption period, period 2 the adoption period and finally, period 3 the post-adoption period. Values are rounded to four decimal places. We test significance using a percentile-bootstrapping procedure. Thus, we get the empirical distribution of the differences of the metric. *, ** and *** indicate significance at the p<10%, p<5% and p<1% level, respectively.

Metric 1.1: Variance of ΔNI* (Panel 1)

			Cluster(s)						
			Absolute values			Differences			
			ERC	SEC	WEC	ERC-SEC	ERC-WEC	SEC-WEC	
Period(s)	Absolute values	1	2.1005	2.2416	1.6589	-0.1411	0.4416**	0.5828***	
		2	1.7196	2.1799	1.5302	-0.4603***	0.1894**	0.6497***	
		3	1.9090	2.3233	1.4453	-0.4143***	0.4637***	0.878***	
	Differences	(1)-(2)	0.3809***	0.0617	0.1286	0.3191***	0.2522***	-0.0669***	Difference-in-differences
		(1)-(3)	0.1915***	0.333	0.2136**	-0.1411***	-0.0221***	0.1190***	
		(2)-(3)	-0.1894	0.2709	0.0850	-0.4603***	-0.2743***	0.1859***	
							Differences-in-differences		

This table presents part of the results of the robustness test where countries with IFRS adoption dates differing from the end of 2005 have been left out of the sample (Israel, New Zealand, Pakistan, Singapore and Turkey). It shows the results of the second earnings management metric - the ratio from equation 3 between the variance of the residuals from a regression of the change in net income on the control variables and the variance of the residuals from a regression of the change in cash flow on the control variables - for each period and cluster (upper-left segment), the differences between clusters for distinct periods (upper-right segment), the differences between periods for distinct clusters (lower-left segment) and the differences-in-differences between periods and clusters (lower-right segment). The differences-in-differences compare the differences between periods over clusters. We number the periods for ease of presentation where period 1 is the pre-adoption period, period 2 the adoption period and finally, period 3 the post-adoption period. Values are rounded to four decimal places. We test significance using a percentile-bootstrapping procedure. Thus, we get the empirical distribution of the differences of the metric. *, ** and *** indicate significance at the p<10%, p<5% and p<1% level, respectively.

Metric 1.2: Variance of ΔNI*/ΔCF* (Panel 2)

			Cluster(s)						
			Absolute values			Differences			
			ERC	SEC	WEC	ERC-SEC	ERC-WEC	SEC-WEC	
Period(s)	Absolute values	1	-0.1104	0.0011	-0.2750	-0.1116***	0.1645***	0.2761***	
		2	-0.1387	-0.0848	-0.3588	-0.0539	0.22***	0.2739***	
		3	-0.1223	-0.1959	-0.3636	0.0736*	0.2414***	0.1677***	
	Differences	(1)-(2)	0.0283	0.086**	0.0838	-0.0577***	-0.0555***	0.0022	Difference-in-differences
		(1)-(3)	0.0118	0.197***	0.0887*	-0.1852***	-0.0768***	0.1083***	
		(2)-(3)	-0.0165	0.1111***	0.005	-0.1275***	-0.0213***	0.1062***	
						Differences-in-differences			

This table presents part of the results of the robustness test where countries with IFRS adoption dates differing from the end of 2005 have been left out of the sample (Israel, New Zealand, Pakistan, Singapore and Turkey). It shows the results of the third earnings management metric - the correlation between the residuals from a regression of cash flows on the control variables from equation 4 and the residuals from a regression of accruals on the control variables from equation 5 - for each period and cluster (upper-left segment), the differences between clusters for distinct periods (upper-right segment), the differences between periods for distinct clusters (lower-left segment) and the differences-in-differences between periods and clusters (lower-right segment). The differences-in-differences compare the difference between periods over clusters. We number the periods for ease of presentation where period 1 is the pre-adoption period, period 2 the adoption period and finally, period 3 the post-adoption period. Values are rounded to four decimal places. We test significance using a percentile-bootstrapping procedure. Thus, we get the empirical distribution of the differences of the metric. *, ** and *** indicate significance at the p<10%, p<5% and p<1% level, respectively.

Metric 1.3: Correlation between accruals and cash flows (Panel 3)

			Cluster(s)					
						Coefficients		
			ERC	SEC	WEC	ERC-SEC	ERC-WEC	SEC-WEC
Period(s)		1				-0.2287*	0.3558***	0.5836***
		2				-0.0991	0.3863***	0.4759***
		3				-0.1856	0.2279**	0.4745***
	Coefficients	(1)-(2)	-0.2183**	-0.0125	-0.2804**			
		(1)-(3)	0.0477	0.101	-0.0504			
		(2)-(3)	0.2855**	0.0989	0.2045*			

This table presents part of the results of the robustness test where countries with IFRS adoption dates differing from the end of 2005 have been left out of the sample (Israel, New Zealand, Pakistan, Singapore and Turkey). It shows the results of the fourth earnings management metric - the coefficient on small positive earnings (SPOS) in the regressions given in equations 6 and 7. In the upper-right segment there are the results of the metric between clusters for distinct periods, in the lower-left segment the results between periods for distinct clusters. Negative values imply that in the cluster comparisons (e.g. ERC-SEC) firms from the first-mentioned cluster (i.e. here ERC firms) are more likely to report small positive earnings and that in the period comparisons (e.g. 1-2) firms are more likely to report small positive earnings in the first-mentioned period (i.e. here in the first period). We number the periods for ease of presentation where period 1 is the pre-adoption period, period 2 the adoption period and finally, period 3 the post-adoption period. Values are rounded to four decimal places. We test significance using a Wald-test, the values for the Wald-statistics are not tabulated. *, ** and *** indicate significance at the p<10%, p<5% and p<1% level, respectively.

Metric 1.4: Small positive earnings (Panel 4)

			Cluster(s)					
						Coefficients		
			ERC	SEC	WEC	ERC-SEC	ERC-WEC	SEC-WEC
Period(s)		1				0.3714***	-0.4342***	-0.7229***
		2				0.7041***	-0.5368***	-1.3058***
		3				0.5237***	-0.4108**	-0.8475***
	Coefficients	(1)-(2)	-0.554***	-0.1913*	-0.7927***			
		(1)-(3)	-0.5089***	-0.3919***	-0.558***			
		(2)-(3)	0.0625	-0.1822***	0.1810			

This table presents part of the results of the robustness test where countries with IFRS adoption dates differing from the end of 2005 have been left out of the sample (Israel, New Zealand, Pakistan, Singapore and Turkey). It shows the results of the fifth earnings management metric - the coefficient on large negative net income (LNNI) in the regressions given in equations 8 and 9. In the upper-right segment there are the results of the metric between clusters for distinct periods, in the lower-left segment the results between periods for distinct clusters. Negative values imply that in the cluster comparisons (e.g. ERC-WEC) firms from the first-mentioned cluster (i.e. here ERC firms) are more likely to recognize large losses and that in the period comparisons (e.g. 1-2) firms are more likely to recognize large losses in the first-mentioned period (i.e. here in the first period). We number the periods for ease of presentation where period 1 is the pre-adoption period, period 2 the adoption period and finally, period 3 the post-adoption period. Values are rounded to four decimal places. We test significance using a Wald-test, the values for the Wald-statistics are not tabulated. *, ** and *** indicate significance at the p<10%, p<5% and p<1% level, respectively.

Metric 1.5: Large Negative Net Income (Panel 5)

		ERC vs. SEC	ERC vs. WEC	SEC vs. WEC
Period(s)	1	-0.3974	-0.4180	-0.4203
	2	-0.3175	-0.2641	-0.3249
	3	-0.5070	-0.4818	-0.4823

This table presents part of the results of the robustness test where countries with IFRS adoption dates differing from the end of 2005 have been left out of the sample (Israel, New Zealand, Pakistan, Singapore and Turkey). It shows the results of the first, return-based accounting comparability metric from equation 23 for the aggregated average accounting comparability between each cluster pair for each period. We number the periods for ease of presentation where period 1 is the pre-adoption period, period 2 the adoption period and finally, period 3 the post-adoption period. Values are rounded to four decimal places. No tests for significant differences have been made due to the long duration of the percentile-bootstrapping calculations.

Metric 3.1: *AGG.ACOMP* (Panel 7)

		Cluster pair		
		ERC vs. SEC	ERC vs. WEC	SEC vs. WEC
Period(s)	1	-0.3228	-0.2322	-0.2408
	2	-0.9070	-0.7511	-0.3468
	3	-0.4955	-0.2022	-0.4432

This table presents part of the results of the robustness test where countries with IFRS adoption dates differing from the end of 2005 have been left out of the sample (Israel, New Zealand, Pakistan, Singapore and Turkey). It shows the results of the second, cash flow-based accounting comparability metric from equation 24 for the aggregated average accounting comparability between each cluster pair for each period. We number the periods for ease of presentation where period 1 is the pre-adoption period, period 2 the adoption period and finally, period 3 the post-adoption period. Values are rounded to four decimal places. No tests for significant differences have been made due to the long duration of the percentile-bootstrapping calculations.

Metric 3.2: *AGG.CFCOMP* (Panel 7)