UNIVERSITY OF TURKU

ABSTRACT

	Bachelor's thesis
X	Master's thesis
	Licentiate's thesis
	Doctoral dissertation

Subject	Accounting and Finance	Date	19.1.2021
Author	Aleksi Somervaara	Number of pages	81
Title	Effects of impression management and ESMA regulation on the quality of non-IFRS earnings reporting		
Supervisor	Ph.D. Antti Fredriksson		

Abstract

This thesis examines the use of alternative performance measures (APMs), more commonly referred to as non-GAAP or non-IFRS figures. These discretionary performance measures that are not based on any common accounting rules have become increasingly popular over the past twenty years. While the reporting entities argue that the APMs may offer valuable information on top of the normal GAAP figures, the adjusted figures also leave room for deception due to the limited regulation governing their use.

The main objective of the thesis is to evaluate the quality of non-IFRS earnings reporting among Finnish listed companies. The empirical tests also analyze whether impression management and the implementation of related ESMA regulation have further effects on the quality. Prior research on the value of non-GAAP reporting is mixed: while the adjusted figures can be backed by both theoretical and practical arguments and evidence, plenty of evidence exists also on misuse of these figures. By considering not only the adjusted figure but also the surrounding communications (through the applied impression management score) this thesis provides comprehensive picture on non-IFRS reporting quality.

In the empirical tests the quality of non-IFRS earnings reporting is analyzed through the persistence of non-IFRS earnings adjustments. The results suggest that from 2012 to 2018 the quality of non-IFRS earnings reporting has been good among the Finnish sample: the adjusted figures are used consistently with the stated definitions. This result could highlight differences in country-specific or time-variant characteristics compared to previous studies since prior results mostly support to the opposite conclusion. The ESMA Guidelines on alternative performance measures are not seen to have any effect on the quality of non-IFRS earnings figures.

When companies are engaging in high levels of impression management the quality of the non-IFRS adjustments drops significantly, which is aligned with previous results. When firms are presenting low-quality adjusted earnings figures they are seen to also employ aggressive impression management to create a misleading view of the performance.

The results are valuable to users of financial information to evaluate the messaging of the companies, as well as for regulators for evaluating the sufficiency of the current regulation.

Key words	Non-IFRS, non-GAAP, alternative performance measure, earnings quality

TURUN YLIOPISTO

TIIVISTELMÄ

	Kandidaatintutkielma
X	Pro gradu -tutkielma
	Lisensiaatintutkielma
	Väitöskirja

Oppiaine	Laskentatoimi ja rahoitus	Päivämäärä	19.1.2021
Tekijä	Aleksi Somervaara	Sivumäärä	81
Otsikko	Vaikutelmien hallinnan ja ESMA regulaation vaikutukset ei-IFRS tulosraportoinnin laatuun		
Ohjaaja	KTT Antti Fredriksson		

Tiivistelmä

Tämä tutkielma tarkastelee vaihtoehtoisten tunnuslukujen eli ei-IFRS tai ei-GAAP tunnuslukujen käyttöä listayhtiöiden raportoinnissa. Näiden tavallisten tilinpäätössääntöjen ulkopuolella olevien tunnuslukujen käyttö on yleistynyt huomattavasti kahdenkymmenen viime vuoden aikana. Ei-IFRS lukuja raportoivat yhtiöt tyypillisesti perustelevat oikaistujen lukujen käyttöä niiden paremmalla informaatioarvolla, mutta paljon huolta on myös kannettu siitä, kuinka näiden niukasti säänneltyjen lukujen avulla on mahdollista vääristellä kuvaa yrityksen suorituskyvystä kauden aikana.

Tämän tutkielman tavoitteena on arvioida suomalaisten listayhtiöiden ei-IFRS raportoinnin laatua. Lisäksi empiirisissä testeissä tarkastellaan ajallisesti otoksen puolivälissä implementoidun ESMA sääntelyn sekä yritysten harjoittaman vaikutelmien hallinnan vaikutusta ei-IFRS raportoinnin laatuun. Aiemmat tutkimustulokset ei-GAAP raportoinnin arvosta ovat osin ristiriitaisia: vaikka näiden oikaistujen lukujen käyttöä tukevat useat teoreettiset sekä käytännölliset argumentit ja tulokset, niin näyttöä on myös siitä, kuinka näitä lukuja väärinkäytetään opportunistisesti. Huomioimalla oikaistuja lukuja ympäröivän viestinnällisen aspektin osana raportoinnin laadun arviointia tämä tutkielma tuo uutta näyttöä ei-IFRS raportoinnin laadusta.

Empiirisissä testeissä ei-IFRS tulosraportointia analysoidaan ei-IFRS tulosoikaisujen pysyvyydellä. Tulokset tukevat ei-IFRS raportoinnin olleen keskimärin laadukasta ja esitettyjen määritelmien mukaista vuosien 2012 ja 2018 välillä. Aiemmat tutkimukset ovat enimmäkseen löytäneet päinvastaisia tuloksia, joten tämä tutkielman eriävät tulokset saattavat liittyä maa- tai aikaperiodikohtaisiin ominaisuuksiin, jotka eroavat aiemmista tutkimuksista. Ajallisesti otoksen puolivälissä implementoitujen ESMA sääntöjen ei nähdä vaikuttaneen ei-IFRS raportoinnin laatuun.

Kun yhtiöt harjoittavat voimakasta vaikutelmien hallintaa ei-IFRS oikaisujen laatu tippuu merkittävästi, joka on linjassa aiempien tutkimusten kanssa. Kun yhtiöt esittävät heikkolaatuisia ei-IFRS tuloslukuja, niin myös näitä lukuja ympäröivä viestintä on aggressiivisempaa, jotta haluttu harhaanjohtava vaikutelma saadaan maksimaalisesti luotua.

Tulokset ovat arvokkaita tilinpäätösinformaation käyttäjille, kun he arvioivat yhtiöiden raportointia, sekä sääntelijöille kun he arvioivat nykyisen ei-IFRS sääntelyn riittävyyttä.

Avainsanat	Ei-IFRS, ei-GAAP, vaihtoehtoiset tunnusluvut, tuloksen laatu

UNIVERSITY OF TURKU
Turku School of Economics

EFFECTS OF IMPRESSION MANAGEMENT AND ESMA REGULATION ON THE QUALITY OF NON-IFRS EARNINGS REPORTING

Master's Thesis
in Accounting and Finance

Author:
Aleksi Somervaara

Supervisor:
Ph.D. Antti Fredriksson

19.1.2021
Turku

The originality of this thesis has been checked in accordance with the University of Turku quality assurance system using the Turnitin OriginalityCheck service.

TABLE OF CONTENTS

LIST OF TABLES

1 INTRODUCTION

1.1 A preface to alternative performance measures and impression management

Alternative performance measure (APM) is a term adopted by the European Securities and Markets Authority (ESMA) for financial measures that are not defined or specified in the relevant financial reporting framework. For European public companies this means that if a certain financial measure is not defined in IFRS, then the measure is considered an APM. Same concept is often also referred to with terms such as adjusted measures, non-IFRS measures, or non-GAAP measures. The non-GAAP prefix can be understood either as a direct reference to the US GAAP or as a non-specific reference to any local accounting ruleset. The latter definition is followed in this thesis, which applies terms APM, non-IFRS and non-GAAP measure mostly interchangeably.[1] An example of an APM is to take IFRS-compliant earnings per share (EPS), subtract transactions that the management considers to be non-recurring, and then present this adjusted EPS in addition to the IFRS EPS. The reasoning for these adjusted figures is that these provide better information to investors, for example by being more comparable between periods or better presenting the performance of the core operations (Entwistle et al. 2012, 232; FINFSA 2019, 10).

The increasing use of these voluntary and historically very lightly regulated measures among public companies is well-documented (e.g. Black & Christensen 2009, 301; Ciesielski & Henry 2017, 35–37; Marques 2017). Looking at 1301 annual reports, Isidro and Marques (2015, 105–106) found that more than 75% of the reports in the sample included at least one APM. More than half of these APMs were based on earnings, of which almost three in four were higher than their normal, GAAP-compliant counterpart. The findings by Isidro and Marques (2015) highlight well the longstanding worry about how these adjusted measures can be used to mislead investors and to manipulate the recipients' impression of the true performance. The possibility for deception is especially high when no or only limited rules are in place for such disclosures. The other side of the argument is that APMs may truly provide incremental, valuable information for investors. With adjusted earnings measures (often also referred to as core earnings, comparable

1 Non-IFRS measure is used only when referring to a strictly IFRS setting, for example in relation to the empirical results. It should be noted that no consensus exists on which term to use in research, and prior studies have also used other terms such as pro forma measures and street earnings or measures.

earnings, or recurring earnings) most commonly the stated purpose of the adjustments is to eliminate the effect of transactions that are deemed non-recurring in nature, or that otherwise are not considered part of the core operations of the firm. This kind of adjusted earnings could be valuable for investors in showcasing the underlying performance of the business without any (so-called) distortions from the accounting standards.

This perspective of contrasting the potential issues with the possible benefits of APMs has been present in the related literature since the very beginning. The first studies were sparked by the increased use of non-GAAP earnings in US companies' financial reports during the 1990s. Bradshaw and Sloan (2002) pointed out how managers and analyst tracking services had started to emphasize non-GAAP earnings measures, which diverged from what the generally accepted accounting principles required. The authors also concluded that investors, too, seemed to prefer these adjusted earnings for their investment decisions. This initial study was hesitant to make definitive conclusions on the relative importance of management opportunism versus altruistic focus on information efficiency as the ultimate factor behind the growing phenomenon. Since these early results, numerous studies related to APMs have been published, providing evidence on the typical characteristics of firms using these figures, what kind of adjustments are the most common, to what extent are APMs used by investors, and more. A thorough review of prior literature is provided in chapter 2.

This thesis continues on the path of Guillamon-Saorin et al. (2017) by linking the evaluation of non-GAAP reporting to impressions management, which can be understood as how firms decide *what* information to present, and *in which manner*, in order to nudge the recipients' impression towards some predetermined goal. The applied impressions management model follows that developed in Brennan et al. (2009) and Guillamon-Saorin et al. (2017). Examining linguistic and psychological tools employed in financial communications is a growing, if not an altogether new trend in accounting literature. Linking together the research on APMs with the analysis of the surrounding financial communications enables a comprehensive examination of how these tools are used: the APMs are evaluated by how these are presented and communicated, not merely by the quantitative factors. The area of linguistics and narratives-focused research within the accounting field is discussed in chapter 2.5, with the focus being on the development of the impression management model applied in this thesis.

1.2 The regulatory background

The regulation of APMs varies internationally. As these measures by definition are something outside of the applicable financial standards, generally they are regulated by the local financial markets authority or similar – if they are regulated at all. In the US, the field is regulated by Regulation G, issued by the Securities and Exchange Commission (SEC) already in 2003. In Australia non-GAAP disclosures are similarly regulated by a regulatory guide issued in 2011. In both of these jurisdictions, presentation of APMs is voluntary, but the applicable regulations must be followed whenever APMs are disclosed. In South Africa non-GAAP measures are regulated, interestingly also partially mandatory. Companies are required to present a specific headline earnings per share, in addition to the IFRS-compliant earnings per share measure (Venter et al. 2014, 2; Marques 2017, 319–323). The development of common European regulation is described well by Isidro and Marques (2015, 97–98), who show how sluggish the process has been. While the predecessor of ESMA issued recommendations on the use of alternative performance measures in 2005, it was only in 2015 that more comprehensive regulation was introduced, in the form of the ESMA Guidelines on alternative performance measures (later also referred to as the ESMA Guidelines or simply as the Guidelines). The ruleset came into force in July 2016.

The goal of the ESMA Guidelines has been to increase the transparency, neutrality and comparability of alternative performance measures (ESMA 2015). Paragraph 4 of the ESMA Guidelines sets that the rules cover APMs reported outside the financial statements, so for example those included in a management report or an earnings release. Additionally, the Guidelines must be followed when measures are presented simultaneously inside and outside the financial statements, so for example both in the financial statements and in the management report (ESMA 2017). In Finland, the Finnish Financial Supervisory Authority (FIN-FSA) is responsible for enforcing the rules set forth in the Guidelines.

The central requirements of the ESMA Guidelines, as defined in the regulation's paragraphs 20 to 48, are presented in Table 1. The rules can be seen to focus on broad principles on how to define and report an APM; almost no direct requirements are set for the content of APMs. Of the five categories of requirements, presentation can be seen as the most important for the objectives of this thesis. In short, this is because it governs the

naming of the measures, meaning that for example a figure labelled as recurring earnings actually must only adjust for non-recurring items.

Table 1 The main requirements of the ESMA Guidelines

PRESENTATION	APMs should be defined in a clear and readable way. APMs should have meaningful labels that reflect their content and basis of calculation: they should not convey misleading messages.
RECONCILIATION	APMs should be reconciled to the most directly reconcilable line item, subtotal or total presented in the financial statements.
EXPLANATION	The use of APMs should be explained, to allow the users to understand their relevance and reliability.
COMPARATIVES	APMs should be accompanied by comparatives for the corresponding previous periods. Comparatives should be reconciled.
CONSISTENCY	The definition and calculation of an APM should be consistent over time. Redefinition or discontinuation of an APM is possible only under exceptional circumstances.

Some notions can be made about the role of IFRS in relation to voluntary disclosures. While the standards do not directly guide the use of APMs, they do set the rules on what is allowed inside the financial statements. The IFRS Conceptual Framework directly states that "… focusing on common information needs does not prevent the reporting entity from including additional information that is most useful to a particular subset of primary users" (IFRS Foundation 2019, A25), leaving flexibility also for voluntary disclosures. The main limitation set by the Conceptual Framework is that the information should be relevant and faithfully represented. Meanwhile IAS 1 *Presentation of Financial Statements* explicitly requires additional disclosures when simply complying with the standards "is insufficient to enable users to understand the impact of particular transactions, other events and conditions on the entity's financial position and financial performance" (IFRS Foundation 2019, A994). The main implication is that there are cases where additional, voluntary information is allowed and even mandatory in IFRS. However, there is little rules for such disclosures' content, apart from the general requirements of The Conceptual Framework. Thus, in Europe the regulation on the use of APMs consist mostly just of the ESMA Guidelines.

1.3 Objectives and scope of the study

This thesis has two main objectives. First, the thesis examines the quality of the adjustments made in calculating adjusted earnings measures. When presented, non-IFRS

earnings in virtually all cases purport to present what is considered the core or the underlying result of the firm. The APMs are described to be derived such that only non-recurring, non-cash, or otherwise non-operative items are eliminated: these transactions should not be of any significance for future performance. By evaluating whether the adjustments have predictive power over future performance, new evidence is provided on whether the non-IFRS earnings figures may be used opportunistically by the reporting firms. If the adjustments reliably predict future performance, then a clear disconnect exists between the stated purpose of the APM and the reality. This is significant also in light of the ESMA Guidelines' requirement on meaningful labelling of APMs.

Similar methodological approach has previously been applied by the likes of Doyle et al. (2003), Lougee and Marquardt (2004), and Leung and Veenman (2018). New contribution over the existing results is provided through several steps. Firstly, a more comprehensive approach is applied by measuring future performance through both earnings and cash flows. Prior studies usually only use one proxy for future performance, but as discussed in Whipple (2015) and later in Black et al. (2018), using both earnings and cash flows allows for better understanding of different justifications for exclusions. Secondly, as is seen in chapter 2, prior literature provides evidence both for the usefulness of the APMs and for these measures being used to exploit investors. New knowledge can be achieved by including additional explanatory variables for understanding the determinants of quality of the adjusted figures. This is done through applying a impression management model, previously seen in Guillamon-Saorin et al. (2017), which is the only prior study to apply a multifactor model of impression management to assess non-GAAP reporting. The model is based on measurement of tone, emphasis, and performance comparisons, thus including both quantitative and qualitative measures of impressions management. This model, of which application is discussed especially in chapter 4.2, allows for examination of whether the opportunistic use of APMs is linked to a more aggressive communicative approach overall. The third incremental contribution is linked to the other main objective of the thesis.

Evaluation of the effects of the ESMA Guidelines on APMs is the second main objective. The rules came into force July 2016, and currently no evaluation on the effects of the regulation on non-IFRS reporting quality has been made. Prior to these guidelines, there was no effective common European regulation on the topic, as has been discussed in other articles previously (Fiechter 2013; Isidro & Marques 2015, 96; Guillamon-Saorin et al. 2017, 452). The implementation of the similar Regulation G in the United States

has been a common research topic, with for example Jennings and Marques (2011) identifying a decrease in the opportunistic use of APMs after the regulation came into effect. Later on some studies have suggested that this effect may have been only temporary (Black et al. 2012; Black & Christensen 2018). This thesis contributes to the existing literature by providing evidence on whether a similar improvement in the quality of adjustments is visible with the ESMA Guidelines. Incremental evidence is also provided on how the impressions management model from Guillamon-Saorin et al. (2017) holds – for better or for worse – in a new setting around the ESMA regulation.

The empirical tests are done on a sample of Finnish public companies' annual earnings press releases[2], which are a mandatory disclosure for the listed companies. These releases normally cover the key figures and tables from the companies' full financial reports, in addition to a CEO comment and summaries of the performance during the period. These are especially valid for the purposes of this thesis because the content represents the summarized main message that the company wants to emphasise. These documents are also more freeform and regulated in lesser amount than the main financial statements. These are also a very common place to present APMs.

The use of Finnish data is supported by several factors. First, while research on European firms' use of APMs is fairly scarce (see Isidro & Marques 2015, 97; Ciesielski & Henry 2017, 40), for Finnish companies this kind of research is practically non-existent, consisting mainly of the informal findings that FIN-FSA has published as part of its newsletter (see FIN-FSA 2017, 2019). For example Ball and Foster (1982) have supported the idea that even replicating existing research in new countries and settings can provide valuable insights. Second, Finland is an interesting setting due to the size of the local stock market. Majority of companies listed on Nasdaq Helsinki (OMXH) are small by international standards, with only a handful of companies exceeding the usual large-cap threshold of ten billion US dollars. This means that the setting enables a sample of what in the global context are mostly mid-cap and small-cap companies. Marques (2017) has previously highlighted how the prior research on APMs focuses on only the largest firms, meaning that little evidence exists on the practice among smaller companies. A third factor supporting the use of Finnish data is the public availability of the annual earnings release data through the officially appointed mechanism (OAM), operated by the stock

[2] The naming is not uniform, but among the sample companies these documents are typically labelled as the *annual financial statements bulletin*, *review,* or *release*.

exchange operator Nasdaq. Through the Finnish OAM a complete and high-quality dataset can be achieved easily, while the public data also allows for easy replication of the results.

Overall, this thesis contributes to the extant literature by providing new evidence on the nature of adjustments made for calculating non-GAAP earnings on a European setting. Prior results are expanded on by combining the analysis of adjustments to an impression management model, which enables a thorough evaluation of the communicative strategy employed with the APMs. Examination is taken even further by providing the first evidence on the effect of the ESMA Guidelines on the quality of adjusted earnings, and by evaluating the applied impressions management model in a new regulatory setting. The results provide valuable findings for the discussion on effects of regulation on the use of non-GAAP figures. The application of the impressions management model from Guillamon-Saorin et al. (2017) offers new validation for the functionality of the model. Additionally, the results provide valuable information for the regulators for evaluating the effects of the ESMA Guidelines, which in its turn may help to improve the quality of information available to all capital market participants and stakeholders.

1.4 Structure of the thesis

Rest of the thesis is structured as follows. The second chapter provides an extensive look into the theoretical background of earnings management, optimal earnings measures, and non-GAAP reporting in general. These topics are tied to the larger trends in accounting research methodology. The chapter concludes with a look into the theoretical background of impressions management, also discussing research on narratives and financial communications.

Based on the theoretical background, the third chapter proceeds with hypothesis development. The fourth chapter details the methodology applied in the thesis through a description of the sample, applied models and construction of the impression management score. After discussing the methods, the results are presented in the fifth chapter, with the sixth chapter providing discussion of the results and the final conclusions. Finally, chapter six summarizes the results and concludes the thesis.

2 THEORETICAL BACKGROUND

2.1 Theories of accounting: the broad perspective

No universal grand theory of accounting exists. This can be seen to be a result of the nature of accounting as a social science and practice: as the saying goes, there is no accounting without accountants. One of the closest thing to consensus that has existed in accounting research may be that the ultimate objective of accounting is to support efficient allocation of capital, as discussed in more length by Kothari et al. (2010). However, this is far from providing any grand theory comparable to the ones in natural sciences. Interestingly Llewelyn (2003) argues that seeking or following such all-encompassing theories may even restrict the field of research in accounting due to the more emergent and local nature of the relevant phenomena. Also, even Watts and Zimmerman (1978, 113), in the seminal paper of positive accounting theory, stated that "management, we believe, plays a central role in the determination of standards", showcasing that the human part of accounting is not completely amiss even in the positive branch of accounting research. While the financial reporting literature is fragmented both in terms of theories and research areas (as also documented by Ball and Foster (1982, 162–163)), the main theoretical developments, as relevant to the thesis at hand, are briefly described below.

Kothari (2001) describes how until the 1960s accounting theories were mainly normative in their nature. Development of theory happened mainly through the researcher's logical reasoning. Implicit in this was that the assumed objectives of accounting were dependent on the researcher, meaning that no common agreement over optimal accounting policies could be reached. Partly motivated by the lack of empiricism, but also by the emergence of the positive economics, efficient markets hypothesis and capital asset pricing model, accounting research started to move towards a more positive approach. Ball and Foster (1982, 171–172) have also highlighted the overall availability econometric techniques as another enabler behind the popularization of the positive research approach. Important was also the introduction of the agency theory by Jensen and Meckling (1976, 308), who defined agency relationship as "a contract under which one or more persons (the principal(s)) engage another person (the agent) to perform some service on their behalf which involves delegating some decision making authority to the agent". Agency theory suggests that when the counterparties are maximizing their own utility there is a good chance that the agent will not act in the best interest of the principal. Therefore, the

principal must incur agency costs, such as monitoring expenditures and bonding expenditures, to better align the incentives of the parties. Agency relationships and costs are seen as a fundamental characteristic of corporations and they are also central to many research topics in accounting, including being relevant for research in non-GAAP reporting.

Two papers by Watts and Zimmerman (1978, 1979) are considered as the starting point for positive accounting theory (later also referred to as PAT), though already some earlier studies had adopted a similar research approach (e.g. Ball & Brown 1968; Beaver 1968). PAT seeks to predict and explain accounting phenomena through empirical methods. The role and the effects of accounting can here be explained through the financial information's ability to minimize agency costs. By minimizing contracting costs companies can increase their chance of survival, providing researchers an optimization problem for predicting accounting choices (Watts & Zimmerman 1990, 133–134).

Explaining accounting choices through the minimization of agency costs represents what is called the efficient contracting perspective, which is one of three major approaches for explaining accounting choices, as identified by Holthausen (1990). Examples of studies supporting this perspective include Whittred (1987) as well as Mian and Smith (1990). The second perspective described by Holthausen (1990) is the opportunistic perspective, which sets that accounting choices can be explained through the management seeking to maximize their personal utility. Important notion is that if it is assumed that management's utility differs from that of the firm's, then this perspective leads to different conclusions than the efficient contracting view. The third approach discussed is the information perspective, which suggests that management chooses accounting methods to provide information about the future cash flows of the firm. As discussed in the article, this perspective assumes that accounting choices provide information about cash flows, but do not actually affect them. The first two perspectives, however, imply that accounting choices have real cash flow effects. Holthausen (1990) proceeds to discuss how the efficient contracting view has gained much less attention than the opportunistic view from researchers, despite the fact that the efficiency perspective could be seen to be more deeply rooted in the history of accounting research and agency theory. Evidence relating to non-GAAP earnings and the efficiency perspective is discussed in chapter 2.3, while findings related to opportunism are described in chapter 2.4. In the broader perspective it can be seen that applying the opportunistic perspective is the more common approach in research on non-GAAP reporting as well.

Researchers' interest in the opportunistic view also highlights the longstanding focus on the need for regulation of accounting, a topic that is of great relevance when discussing non-GAAP reporting and whether it should be more strictly regulated. A general argument for regulating corporate financial disclosures is that without regulation the information markets will fail. This could happen by two main ways: due to the nature of financial information as a public good, or due to the information asymmetry inherent to capital markets. It is not, however, clear whether the benefits of regulation exceed the related costs, and any ultimate conclusions on the optimal level of regulation have not (or even cannot) been made. (Cooper & Keim 1983) It is mostly uncontroversial that regulation is warranted when it is able to correct market failures, such as problems with information asymmetry (differences of opinion more commonly exist on whether the regulation is actually able to lessen the impact of market failures). Chapters 2.3 and 2.4 discuss the research on non-GAAP reporting at more depth, but the existing literature does offer evidence of APMs being used opportunistically and contrary to the publicly stated objectives – suggesting a market failure. This thesis adopts the view that in such cases regulation of financial reporting is warranted.

It is important to consider the pitfalls of positive accounting theory, which has faced lots of criticism as well. For example Whitley (1988) early on suggested that PAT fails to consider the appropriateness of positive methods for social sciences. The idea that any grand theory of social sciences, and especially one based on equilibrium economics, could predict everything was met with great skepticism. Further, the article found PAT internally contradictory because while stating to follow empiricism and its ideals, the theory adopted concepts such as efficient markets hypothesis and CAPM, which were seen to fail all empirical tests of validity. Meanwhile Chambers (1993) took issue with the fact that PAT relied too much on capital markets and failed to design experiments with proper controls. The author argued that accounting research is not the same as research in economics, and thus should not focus just on market reactions, EMH and CAPM. Tinker et al. (1982) have presented the idea that while PAT tries to claim validity by being interested only in facts and observations, it also is value-laden in its own way. The argument being that it is a question of faith and values to adopt an idea of efficient markets or that information efficiency is a critical factor for wellbeing, and that the theory tries to hide these kinds of implicit values. Christenson (1983) even took problem with the adoption of the term "positive", suggesting that "empirical" would be more suitable. It was seen that a positive theory that allows for exceptions could not be considered valid: while a

predictive theory can be useful, it certainly does not present positivism, in the form that the term is generally applied in philosophy of science.

Despite the criticism, positive accounting theory and its characteristics continue to reflect the philosophy that a significant portion of accounting studies adopt, even if the field has somewhat moved on from PAT as a term. The criticism did spark response from the proponents of PAT. For example Watts & Zimmerman (1990) denounce some of the criticism for not considering the growing body of evidence in favor of PAT. They also discuss the criticism relating to research methodology, agreeing that issues such as omitted variables and alternative hypotheses can lead to problems in conclusions. This, however, was seen as a matter of refining the models and explanations; the suggestion being that if and when better models are discovered, the positive accounting community is ready to welcome these as the field develops. Watts & Zimmerman (1990) also later agreed that positive research is indeed value-laden, but argued that this is not of great importance, because the usefulness of a positive theory is derived from its ability to predict and explain a phenomenon. The implicit values are only relevant when they interfere with the predictive or explanatory ability of the theory. Arguably the balanced view is that both prescriptive and positive research in accounting have their place. For example Kaplan and Ruland (1991, 362) argue that "it is in the public interest that both positive and normative theories of accounting be developed." This kind of view is not uncommon (e.g. Watts 1977; Jensen 1983; Beattie 2014). As Watts (1977, 57) describes, "the development of prescriptions and the development of theory are not incompatible. The development of prescriptions which are likely to achieve their objectives requires an underlying theory which explains observed phenomena: which predicts the effects of particular prescriptions."

In light of the above discussion, this thesis adopts many of the characteristics linked to PAT and empiricism in general, similar to most of the prior literature on non-GAAP reporting. In studying the disclosures and communications of the reporting firms, the empirical results of this thesis provide evidence especially about the opportunistic perspective of positive accounting. While this thesis does not, or need to, strictly follow the ideals of positive economics and the efficient markets hypothesis, merit is seen in prior studies adopting these perspectives, and results from many such studies are discussed in following chapters. While the focus of the thesis is on providing new empirical evidence on the subject matter, the prescriptive view is not ignored. A main objective of this thesis is to provide valuable information for regulators and users of financial statements on how to approach the use of alternative performance measures and impressions management.

2.2 The importance of earnings and the motive for non-GAAP disclosures

Before delving into the micro-level perspective on the use, effects, and implications of alternative performance measures, it is important to consider the reasons behind why these adjusted figures are reported in the first place. From the reporting entities' perspective, it is most often stated that earnings-related APMs are reported in order to present the core or the recurring results of the business operations. This means that the effects of non-recurring or non-cash transactions are adjusted out of the GAAP figure. For example when looking at Finnish companies, FIN-FSA (2017, 10) found that most commonly it is stated that either the APM is considered more comparable between reporting periods, or that it gives a better understanding of the operational business development. This is in agreement with what the academic literature shows as well (Black & Christensen 2009; Shiah-Hou & Teng 2016; Black et al. 2018). Implicit in this is that earnings are seen as an important measure of performance (hence the focus on them), but that for some reason the GAAP version is considered insufficient or at least suboptimal.

Skepticism towards accounting earnings and their usefulness is not a new phenomenon within the research community. Kothari (2001, 113) documents how in the mid-twentieth century there was increasing disagreement on the usefulness of accounting income. This was considered to be stemming from the lack of generally agreed objectives of financial accounting. This skepticism led to two influential studies in the late 60's. Ball and Brown (1968) conducted what is one of the earliest and most influential capital markets studies by examining how the markets react to unexpected changes in income. The results of the study supported the idea of accounting income containing meaningful information for the capital markets. Beaver (1968) was similarly interested in the information content of earnings, meaning whether the earnings reports lead to adjustments in the expected future returns. Using return volatility during the earnings announcement period (as opposed to unexpected earnings changes) Beaver (1968) found positive support for information content of accounting earnings. Some recent studies such as Beaver et al. (2018) have presented similar results.

The reason for why accounting earnings are taken to contain valuable information could relate to the figure acting as an important top-level performance gauge. For example when surveying CFOs, Graham et al. (2005) found that managers see earnings as the ultimate metric reported to external stakeholders. Interesting finding was also that non-GAAP earnings were seen to be as important of a metric as cash flows, suggesting that at

least for external reporting cash flows can be deemed a secondary metric compared to earnings. This is despite the often-prevalent focus on cash flows. An intuitive logic for preferring cash flows over earnings is the so-called realness of cash: cash flows are a concrete measure of how the business creates or consumes cash, as opposed to earnings which ultimately is a pure accounting concept. This is also seen in basic valuation methods such as discounted cash flow models where, as the name suggest, it is theoretically correct to discount future cash flows, not future earnings.

There is, however, empirical support for considering earnings over cash flows as the key performance metric of a firm – even within the aforementioned context. Dechow (1994) reviewed previous literature and pointed out that while both earnings and cash flow were seen to offer incremental information on top of each other, no conclusion had been made on which of the two serves as the better summary of firm performance individually. Correspondingly the study sought to understand which of the two serves as the better measure of performance separate from the other one. The conclusion was that cash flows can be seen as the better measure only over much longer time periods, because on a short and finite timescale the cash flows contain too much noise. Over finite time periods accruals and matching help earnings act as the more valuable measure of performance. A related finding is presented by Barth et al. (2001), who describe how individual accruals components can have meaningful information about future cash flows. A simple explanation for preferring earnings is that while theoretical valuation is concerned with future cash flows, when comparing accounting earnings and accounting cash flows (which represent historical information), it is earnings that is the better predictor of future cash flows (Kim & Kross 2005; Jordan et al. 2011).

The importance of earnings is also highlighted by the extensive literature on earnings quality (EQ) and earnings management (EM). Both of the terms lack universal definitions, but for example Dechow et al. (2010, 344) define earnings quality in that “higher quality earnings provide more information about the features of a firm’s financial performance that are relevant to a specific decision made by a specific decision-maker”, meaning that EQ can be seen as completely context dependent concept. Dechow et al. (2010) further identify three categories of EQ measures: properties of earnings (including earnings persistence and accruals), investor responsiveness to earnings, and external indicators of earnings misstatements (including restatements and publicized internal control deficiencies).

Earnings quality can also be examined by deriving it from earnings management, which has a decreasing effect on earnings quality. In broad terms EM could be defined as the use of discretion in accounting, reporting, and real economic decisions to manipulate the earnings figure towards some specific goal. EM as a research topic got more popular after the turn of the century, and since then the field of research has broadened further. (Walker 2013, 445–446) Three main categories of earnings management practices are seen to exist: accruals earnings management (AEM), real earnings management (REM) and classification shifting. Of these AEM was the dominant topic of research especially up until early 2000s. AEM refers to discretionarily choosing to present accruals under certain year to achieve the wanted earnings figure. It reflects a pure accounting adjustment and may also constitute as a breach of accounting law if the accrual basis is not followed correctly. REM refers to influencing the earnings through concrete business decisions. For example, when appearing to be missing the earnings estimates, a firm might decide to postpone all R&D operations to the following year, in order to avoid having the related costs lower the earnings of the current period. While this may not always constitute as fraudulent behavior, it can still be against the long-term interest of the shareholders. Finally, classification shifting is about shifting items between income statement rows, for example moving expenses from operating expenses to non-recurring expenses or to below operating profit subtotal. While the income statement total (net profit) remains unchanged, the presentation differs and can give a misleading picture about the firm's performance. (Abernathy et al. 2014, 603–605)

Especially classification shifting and concerns related to it are very similar to the potential issues with non-GAAP earnings. In both cases the potential problems relate to management discretionally adjusting reported figures to mislead investors. When adopting a broad definition of earnings management, reporting non-GAAP earnings measures could be seen as a fourth category of EM practices. For example Walker (2013) and Fan et al. (2010) indeed see non-GAAP earnings reporting as a part of the earnings management field. This is justifiable especially when considering the goals of more traditional EM techniques and non-GAAP reporting. With both practices the objectives relate to achieving earnings-based contractual obligations or targets, as well as to influencing the information set used to evaluate the firm (Walker 2013, 457).

Discussing the trade-offs between the different EM methods, Abernathy et al. (2014) describe how firms' move from AEM to REM when the ability to engage in AEM is constrained, and from REM to classification shifting when the use of REM is also

constrained. Including non-GAAP reporting as additional facet of earnings management could improve this framework further. Considering the increased awareness and scrutiny of auditors and regulators on the more longstanding EM methods such as accruals management, non-GAAP reporting can act as a cost effective and secure option for managing the perceptions of the audience. Completely neglecting APMs as part of the earnings management toolset used by the management offers a too restrictive picture, even if the APMs may also help to improve the information environment. This thesis supports the view that at least when adopting a broad definition of earnings management, non-GAAP reporting should be seen as a subcategory of EM.

The goal of non-GAAP earnings reporting or earnings management is not to simply maximize the reported earnings. Roughly until 1980s most of the earnings management literature focused on income smoothing, which today is seen as one subcategory of earnings management (Walker 2013, 447). This relates to companies purposefully decreasing the variance of their earnings, to make the earnings appear more persistent and predictable. This is desirable since according to both the classic valuation models and also simple behavioral evaluation, more variable earnings imply riskier business. This is also verified in the study by Graham et al. (2005, 44–47), where surveyed CFOs overwhelmingly supported the idea that they prefer smooth earnings, not highly volatile ones. In the relevant literature this can be understood through the work on earnings response coefficients. As Kothari (2001) summarizes, together with risk, growth and interest rate, the persistence of the earnings is one of the main determinants for the market reaction on an earnings release. If the unexpected change in earnings is expected to persist and present the new normal, then the change leads to a stronger market reaction. Simple one-off gains warrant only much smaller earnings responses – if any at all. Taking this into consideration, it is easy to see how smoother earnings make it simpler for the market participants to evaluate the future earnings as well.

Other research supports presenting smooth and predictable earnings as well. Brown and Caylor (2005) found that after the 1990s managers seem to avoid negative earnings surprises rather than avoiding negative or declined earnings in general. This was seen to be a result of increased media coverage, where missing the expected or forecasted earnings figures led to more significant declines in the stock price than when just having the earnings decline in absolute terms. Jorgensen et al. (2012) also found that higher earnings dispersion leads to investors demanding higher expected returns, thus pushing the share price down. According to Brown et al. (2015, 4) analysts want to see sustainable and

repeatable earnings, that should also reflect the economic reality and be supported by operating cash flows. It may be that analysts and investors reward smooth earnings because they intuitively appear to be sustainable and repeatable, even if the low variability is a result of earnings management and not reflecting the economic reality.

Related to what kind of earnings analysts prefer, it is important to note that analysts nevertheless tend to make their own adjustments to reported accounting figures. For example, analyst may eliminate some one-time or non-cash transactions from the IFRS earnings, similarly to what reporting firms do in deriving the APMs. However, no consensus exists on what should be eliminated from IFRS-earnings to achieve a so-called perfect earnings number. Different analysts apply different practices. (CFA Society UK 2015) What can be said is that analysts adjust the result figures so that the adjusted earnings figure would match what was described by Brown et al. (2015, 4): that the figure could be considered comparable and reflecting of economic reality.

Overall, earnings can be deemed one of the most important figures that a company provides to its stakeholders. Further, investors and analysts have been shown to prefer certain type of earnings, hence management has been seen to engage even in non-value-adding activities to fulfill the expectations of the stakeholders. These expectations and impressions can also be affected through non-GAAP earnings reporting. As seen above, the arguments for disclosing earnings-related APMs are reflective of what stakeholders have been shown to seek and prefer in performance measures. Next the examination moves to the effects and usefulness of non-GAAP earnings, and after that to what extent the reporting might also be motivated by management opportunism.

2.3 The value of non-GAAP earnings: the efficiency perspective

Whether alternative performance measures provide incremental value for different stakeholders of a company has been a popular topic of research. This line of work can be seen to present the efficiency perspective, which seeks to explain accounting choices through minimization of agency costs. As Holthausen (1990) discusses, very strict tests of efficiency-based hypotheses are difficult and relatively rare, and not all of the studies presented in this subchapter form their hypotheses strictly through efficient contracting. The literature discussed here does however focus on arguments on effects, usefulness, and value of non-GAAP reporting, and these factors are concretely linked with the efficiency view. If the evidence suggests that the APMs provide valuable information and improve the information environment, then this would support the efficiency perspective of

accounting method choice. As seen in the previous chapter, there is theoretical support for why APMs could have effects of this kind. Investors are interested in the firms' capability to produce sustainable and repeatable earnings, and professional analyst ultimately adjust reported GAAP figures in any case. To what extent the literature supports the efficiency view with APMs is discussed next.

Already the first studies on non-GAAP reporting provided some initial evidence on the usefulness of these measures. The specific approach of these studies was to evaluate whether the disclosed APMs affect investor decisions and the valuation of the company. Bradshaw and Sloan (2002) were interested in the growing phenomenon of adjusted earnings measures, which they referred to as street earnings. They documented how the prevalence of these measures had grown immensely between 1985 and 1997. By looking at the association between long window stock returns and both GAAP and non-GAAP earnings, the study found that in the beginning years of the sample, there was no significant difference between the reactions to the two earnings measures. However, coming into the 1990's, the value relevance of normal GAAP earnings had declined notably, while for the non-GAAP figures the value relevance had increased. Investors seemed to prefer adjusted non-GAAP earnings to the figures reported strictly in accordance with the local accounting laws and regulations. Similarly, Bhattacharya et al. (2003, 288) also found non-GAAP earnings to be "significantly more informative to investors than GAAP operating earnings" and that they were seen "a more permanent measure of firm performance". Here the conclusions on informativeness were based on short-window cumulative abnormal earnings regression model. Over the applied three-day event window, the earnings response coefficient for non-GAAP earnings was significantly positive, while for ordinary GAAP earnings the coefficient did not statistically differ from zero. Lougee and Marquardt (2004) utilized a similar regression model over a two-day event window, and the results again suggested that the adjusted earnings had the higher information content. This article also provided additional findings related to firms reporting adjusted earnings, including that firms reporting these measures on average were characterized by greater sales growth and earnings variability, and also that the earnings response coefficient of these firms' GAAP earnings tended to be relatively lower. Overall, the literature on the association between non-GAAP disclosures and capital market reactions tend to agree across the board that a significant effect is in place: investors do find adjusted figures useful and use these for their investment decisions.

The research on informativeness and usefulness of adjusted figures is not limited to only looking at stock market returns, which as an approach has its detractors. Before moving on completely from capital markets considerations, there is still one research approach tied to capital markets, but not to stock prices directly. Christensen et al. (2014) examined usefulness of non-GAAP disclosures through short selling activity. This allowed the researchers to analyze to what extent non-GAAP disclosures seem to provide tangible information for the more sophisticated investors. The main finding was that short selling was significantly more active when the earnings report contained an APM, as opposed to earnings reports that did not contain one. Through additional analysis the authors also concluded that short sellers were particularly active when the adjusted measure was achieved through removing recurring items from the GAAP figure. It was also found that highly shorted stocks in this context were associated with significantly negative abnormal returns over a five to twenty-day time period. The findings suggested that the adjusted figures have valuable information at least for the sophisticated investors. Considering this, it could be seen worrying that for example Bhattacharya et al. (2007) have shown that during the earnings announcement day it is the less sophisticated investors who are reacting most actively to the disclosed non-GAAP information.

Moving on from pure capital markets approaches, the most prevalent research method has been to assess the predictive ability of non-GAAP earnings, either in absolute or relative terms. As already discussed, most commonly the argument for presenting adjusted earnings is that the APM better presents the recurring result of the firm. Meaning that the measure is derived, for example, by eliminating non-recurring and non-cash transactions out of the GAAP figure. Because the adjusted figure is implied to present the recurring result, the adjustments themselves should not have any predictive ability over future performance. In this case the APM would provide valuable information, as the reporting firms suggest. In an early study, Doyle et al. (2003) modelled non-GAAP earnings' ability to predict future cash flows. This was reasoned to be a valid method as the sample firms contended that they were only adjusting for non-cash items in calculating the APM. The authors found that over a three-year period one US dollar of excluded items predicted over three US dollars of negative cash flow in the future: items that were stated to be non-cash were anything but. The excluded items were further divided into two groups. Adjustments that could be readily classified as one-offs were labelled "special items", while the remaining, more vague adjustments were labelled "other exclusions". Upon further analysis, it was seen that the special items were mainly irrelevant in

predicting future cash flows; the predictive ability was stemming from the vaguer other exclusions. Finally, the authors also considered if this is reflected in the share price, concluding that on a near-term period of three days the price declines accordingly, but over a three-year period there is a clear mispricing of the stock. The implication is that while investors initially react to the disclosed APM properly, they are misestimating the future cash flow implications. Therefore, the study concluded that while investors use the non-GAAP measure as it were useful, the measure actually misleads their investment decision.

Lougee and Marquardt (2004) examined the non-GAAP earnings' ability to predict future earnings. Their conclusion was that the adjusted figures did not have predictive power over future GAAP earnings, but actually had some predictive power for future non-GAAP earnings. This suggesting that the adjusted figure could at least be considered more internally persistent than the GAAP figure. On a more recent study Leung and Veenman (2018) documented that in loss-making firms the GAAP earnings have very little predictive power over future performance, but non-GAAP earnings had very high predictive power. The future performance was here operationalized through both cash flows and earnings, a method that is aligned with the discussion in Whipple (2015) and Black et al. (2018). In one of the more comprehensive comparisons between non-GAAP and GAAP earnings, Ribeiro et al. (2019) compared these figures over a 14 year period for 500 companies. Here the conclusion was that the non-GAAP figure was superior in almost every situation: they were more persistent and had more predictive power over future earnings. When considering the non-GAAP measures' predictive power for future earnings, as compared to GAAP measures, the literature suggests that the adjusted figure has at least limited edge over the latter. With predictive power over cash flows the evidence is somewhat more mixed. This might be explained by the fact that for example the study by Doyle et al. (2003) used a sample from an era where the Regulation G was not yet effective. The use of APMs was less restricted and may have been more opportunistic at this point in time. The studies have also differed in over what term the predictive power has been considered, as Doyle et al. (2003) considered predictive power over a three-year period, while often a more immediate model is used.

Direct survey responses could also be used for determining if APMs are useful. CFA Society UK (2015) provides insight about how analysts feel about APMs and how these professionals use these measures. In the survey 61% of analysts stated that they used the adjusted earnings figures reported by management. The voluntary disclosures were thus generally seen beneficial, even though respondents also directed attention to the fact that

they rarely took the APMs at face value. This was also reflected in the fact that 60% of analysts felt that the IFRS figures are more reliable than the discretionary ones. The general conclusion was that APMs were especially useful when and in that they provided information on what management sees as non-recurring, allowing the analysts to do their own conclusions based on this. This also reflects the study by Elliott (2006), which showed how reconciliating the APMs to the most directly comparable GAAP item increases the reliability of the measure. The reconciliations can help by better enabling the users of the financial statements to understand the adjustments and evaluate if these seem sensible. The ESMA Guidelines correspondingly require APMs to be reconciled to the closest IFRS line item. Other surveys in the series provide additional insights. Analysts seem to use APMs both as a direct and indirect valuation input, and also as a measure of accounting quality (CFA Institute 2016). Similarly interestingly, analysts seem to support expanded use of APMs, while simultaneously pressing for more formal regulation of these (CFA Institute 2018).

Other, more novel approaches have also been employed to identify how certain factors may affect the value of the APMs. Isidro and Marques (2015) examined what effect local corporate governance factors such as investor protection and efficient law enforcement had for voluntary disclosures. The results showed that higher quality corporate governance was associated with decreased usefulness of voluntary disclosures. The explanation for this observation was that efficient enforcement of local laws and regulations makes it next to impossible to manipulate the GAAP earnings. In such environment presenting misleading APMs may be the firm's only option to try to discretionarily affect the recipients view of the performance of the firm. This can be contrasted to the findings of Abernathy et al. (2014) on how firms use different earnings management tools depending on which options they have available, as was discussed in chapter 2.2. Similarly interesting is the study by Guillamon-Saorin et al. (2017), which as discussed in the introduction, serves as the background for this thesis as well, and is discussed more in chapters 2.5 and 4.2. It is worthwhile to briefly point out that the article did provide evidence on how non-GAAP earnings do generally provide useful incremental information to the markets, though when these figures were linked with high levels of impressions management the investors actually penalized the firm for this. This would suggest that investors do apply APMs but may also critically evaluate and apply skepticism when interpreting them.

The literature overall is quite clear in that non-GAAP disclosures do provide incremental information above the mandatory figures and that investors certainly act on this

information; similar results have been achieved through many different methods and samples. Evidence suggests that non-GAAP earnings have the better predictive ability for future earnings, though some more mixed evidence exists as well. Plenty of support exists for the efficiency view: these voluntary disclosures may in large part exist to improve the information environment. However, as these figures have been shown to lead to identifiable market reactions, it is important to also consider the other side of the coin: to what extent the APMs may still be distorted by opportunistic or mischievous motives.

2.4 Opportunistic use of non-GAAP reporting: the opportunistic perspective

Lots of attention has been paid to the potential misuse of APMs, reflecting the opportunistic perspective discussed in chapter 2.1. In addition to the media and the academic field, also the regulators and standard setters have become increasingly active regarding the topic. For example, the IASB chair Hans Hoogervorst has expressed his concerns on the use of APMs in two influential speeches. In the first speech Hoogervorst (31.3.2015) paid attention to how APMs give only a "selective presentation of an entity's financial performance", which according to him often "is not free from bias". A year later Hoogervorst (11.5.2016) described his concern over the fact that over 70% of companies in the US had been shown to try to present APMs that give a too favorable picture of the firm's result. With these concerns in mind, what does academic research have to say about opportunistic use of APMs?

By and large there are two factors that may lead managers to use APMs opportunistically. The first factor is that in order to maximize the enterprise value or the stock price, firms would need to meet or beat the analyst estimates (Kasznik & McNichols 2002; Matsumoto 2002; Brown & Higgins 2005). If the analyst estimates are not reached through normal business activities or through other earnings management practices, companies may try to present misleading APMs that meet or beat the consensus estimates. The second factor relates to the fact that the compensation of the top management tends to be very bonus-laden, meaning that the compensation is directly tied to the stock price. Numerous earnings management studies have provided evidence that management may and often do act opportunistically in order to maximize their own compensation (e.g. Richardson et al. 2004; Brockman et al. 2010).

The simplest and most common way to examine this topic has been to look at whether the non-GAAP figures routinely exceed the GAAP counterpart. If the discretionary figure is always more positive than the figure that is based on a common ruleset, then this could

imply opportunistic use. Majority of the studies show that it indeed is the case that APMs give a more favorable view than the closest GAAP-compliant measure. FIN-FSA (2017) has shown that this is the case in Finland, finding that in their sample only one of the adjusted earnings metrics was lower than the IFRS compliant metric. Isidro and Marques (2013) have similarly presented that the non-GAAP figure exceeds the GAAP figure in majority of the cases, with the statement being based both on their own empirical research and review of previous studies from the US. Later Isidro and Marques (2015, 96) found the voluntary earnings measures to exceed the GAAP figure 72% of the time in their European sample. Similar results can be seen among the Australian companies' as well (Cameron et al. 2012).

This kind of analysis is of course fairly simplistic, and the fact that the APM tends to exceed the GAAP counterpart is not definitive proof of opportunistic reporting. One reason for this is that accounting in general is conservative in its nature. Even with IFRS, where capital markets are considered to be on the forefront and traditional conservatism is less pronounced than for example in the US GAAP, this is still visible in numerous places. Examples include inventory valuation (measured at the lower of cost and net realizable value), revenue recognition (allowed when the performance obligation is fulfilled; for example a firm order and capability to fulfill said order is not enough), asymmetric recognition of contingent assets and liabilities (former cannot be recognized) and impairments of goodwill (cannot ever be reversed despite supporting evidence). The latest update to the IFRS Conceptual Framework also explicitly added prudence as a term to the IFRS lexicon. While this may be beneficial from the stewardship point of view, arguably it does not support the information needs of the investors, who are more interested in exact values that reflect the market prices. A reasonable explanation for the APM exceeding the GAAP figure could be that firms are at least partially undoing the effects of GAAP distortions such as conservatism or prudence. Since conservative accounting understates the financial performance and financial position, this could lead to the adjusted figure persistently exceeding its GAAP counterpart. This kind of view would be aligned with the findings of Dichev and Tang (2008), who pointed out how moving toward fair value accounting may be necessary to provide earnings measures that are actually relevant for the financial markets.

Accordingly, other methods of identifying potential misuse of APMs have been devised by researchers. Another simple method is to consider the reporting frequency of these measures. The hypothesis is that consistent use and disclosure of APMs is indicative

of acceptable reporting practice. Only sporadic use of APMs would imply opportunistic behavior, since this could imply that these are used only when it is necessary to boost the recipients' perception of the earnings. This kind of examination is justifiable especially since the lack of consistency and comparability between APMs can be considered their main problem (Ciesielski & Henry 2017). The ESMA Guidelines also directly require that the APMs should be defined and calculated consistently over time; redefinition or discontinuation is allowed only under exceptional circumstances. Black and Christensen (2009) concluded that when APMs are used in a continuous manner, the use indeed is less aggressive. Curtis et al. (2013) however showed that actually less than half of the firms in their sample operated completely consistently in reporting non-GAAP figures. Similarly interesting is how Isidro and Marques (2015, 105–106) found that of the firms reporting non-GAAP earnings in their sample during 2003 to 2007, only 30% actually did this each year. This indicates that APMs might be used only when necessary for achieving the wanted impression of the result, for example when the firm is missing certain earnings benchmark. These results could be defended from the reporting entities' point of view through the firms not having any so-called non-core transactions every year: there simply was no need for any adjustments. Thus, even more thorough evaluation of reporting practices is needed for more definitive conclusions.

The most thorough analysis is achieved when the actual adjustments made in deriving the APM from the GAAP figure are analyzed, meaning that attention is paid to what exactly is taken out of (or added into) the original figure. This kind of analysis is difficult and always at least partially relies on subjective decisions by the researcher. APMs as reported by the company are not readily available from any common database, which means that the gathering a proper sample involves lots of slow manual work. Because of this, a common methodological approach has been to utilize measures such as I/B/E/S adjusted earnings as a proxy for reported APMs, a method of which validity is discussed in chapter 4.1. Gathering information on individual adjustments made to reach the APM is an even more involved task, not least because detailed information is not always available at all. Some studies have still attempted to do this kind of item-level evaluation.

Commonly the adjustment items are categorized as being either normal, recurring items or one-time, non-recurring items. This enables researchers to compare the practices of different firms. It is also valid in the sense that analysts generally exclude one-time items from their earnings forecasts as well (Brown et al. 2015, 25). Black and Christensen (2009) provided evidence that while the stated objective of APMs was to provide the

result without non-recurring items, meaning that the figure should tell about the performance without any one-time items, companies continuously adjusted lots of recurring items out of the non-GAAP measures. The most common exclusions were seen to be items related to depreciation, research and development and stock-based compensation; none of which are patently non-recurring by their nature. Research and development expenses are commonly described as necessary for succeeding in the marketplace and can hardly be deemed to be outside of the vital operations of most firms. Stock-based compensation also most often is a constant part of the overall compensation package of top management. Labelling it as completely non-recurring item would be at least a suspect practice. In regard to depreciation-related items, eliminating them from the earnings is suspect in the sense that analysts generally do include these items in their calculations (Brown et al. 2015, 25). Similar results about companies commonly excluding research and development, depreciation and stock-based compensation from the APMs were identified by Isidro and Marques (2015) as well.

In a recent study Black et al. (2018) found that over the past ten years companies have increasingly adjusted recurring items out of the non-GAAP figures. The most interesting finding of the study relates to comparing different types of adjustment items by their frequency and variation. Frequency presents how often the exclusions is made and variation how much the value of the item changes between the times it is adjusted out of the APM. The study found that while stock-based compensation and amortization were commonly labelled as non-recurring items and adjusted out of the non-GAAP figure, they actually had a high frequency score: these recurred almost every year in the sample period. These items also had low variation scores, which implies that they were continuously reoccurring in rather similar amounts. These results indicate rather strongly that these items are not non-recurring in their nature. Several other studies have shown companies adjusting recurring items out of their non-GAAP earnings as well (e.g. Isidro & Marques 2013; Shiah-Hou & Teng 2016). The general consensus from prior research suggests that companies do commonly exclude recurring items from their discretionary earnings figures, even when stating that the measure is derived by eliminating only the effects of non-recurring transactions.

Another way to obtain evidence about potentially mischievous use of APMs is to consider external factors as well. As referred to earlier, meeting or beating the earnings benchmarks or forecasts has been shown to be a major driver of the stock price. Accordingly, several studies have looked at if APMs are used as an earnings management

practice to reach these benchmarks. Isidro and Marques (2015) showed that presenting APMs is more common when the companies are missing relevant benchmarks, while Doyle et al. (2013) provided evidence that reporting non-GAAP figures is linked to exceeding analyst forecasts. Michael (2019) found that among Australian companies the adjusted earnings on average exceed the analyst forecasts. This kind of earnings management to meet or beat the analyst estimates is more common in countries with high level of investor protection, because in these countries relatively few alternatives exist for tampering with the results (Brown & Higgins 2005). This can be seen worrying in at least two ways. Firstly, the exclusions made by management are generally of lower quality than the adjustments made by analysts (Bentley et al. 2018). Secondly, meeting or beating the consensus estimates earns a market premium independent of the actual absolute performance, even when there is evidence that the result was achieved through earnings management (Bartov et al. 2002).

Compensation and corporate governance are other relevant external factors to consider. In the earnings management literature it has been shown that there is a significant increase in the frequency and magnitude of positive corporate announcements in the pre-exercise period of management stock options when CEOs implement exercise-and-sell strategies (Brockman et al. 2010). As for studies considering non-GAAP information directly, Shiah-Hou and Teng (2016) found that non-GAAP information is more common when the CEO or the CFO sells shares 2 weeks after the earnings release. Bansal et al. (2013) on the other hand found that when managers have more of their compensation tied to stock volatility sensitivity, then the frequency of reporting APMs is higher. Based on this they concluded that compensation arrangements do have their role in promoting good reporting practices. This is true with the compensation of the board of directors as well, as it has been shown that when this is contingent on market performance, providing non-GAAP information is more likely (Isidro & Marques 2013). On the other hand, efficient governance and independent board members is linked to less opportunistic use of APMs, and these two have also been shown to increase the APMs' predictive power over future earnings (Frankel et al. 2011; Isidro & Marques 2013).

Overall, plenty of evidence exists on non-GAAP reporting being at least partially motivated by opportunistic factors. This conclusion is even further supported by the fact that implementing new regulation has been shown to lessen the issues mentioned above, at least in the initial years after the regulation (Marques 2006; Jennings & Marques 2011; Black et al. 2015; Bond et al. 2017; Black & Christensen 2018). The introduction of the

ESMA Guidelines on APMs and the fact that the IASB is considering their own guidance on APMs and income statement subtotals can therefore be seen as welcome developments, which may lessen the negative effects stemming from non-GAAP reporting. In the big picture prior literature suggests that non-GAAP disclosures are motived both by market efficiency-related and opportunistic motives, as has previously been discussed by Leung and Veenman (2018). This means that further studies on APMs need to move beyond the more simplistic models and include other variables for explaining when the reporting might be more reflective of efficiency versus opportunism. In this thesis this is done by taking into account the wider communicative perspective and impressions management, of which theoretical background is discussed next.

2.5 Narratives and linguistic methods in accounting research

Especially when evaluating the opportunistic use of non-GAAP figures, it is easy to see why the context of communication matters. Under such assumptions one would expect that adjusted figures are used similarly to classification shifting, which was discussed in chapter 2.2: the goal is to mislead and to obfuscate the recipients' view of the actual performance. Such goal is not achieved through the figure in itself, but only through a corresponding adjustment in the firm's overall communications, meaning how these adjusted figures are then actually presented. This is also why the ESMA Guidelines focus especially on presentation technical matters, such as how the adjusted figure must not be emphasized over its regulation-based counterpart.

Beattie (2014) offers an impressive summary of research on narratives within the accounting field, covering both the positivist and the interpretative lines of research. The narrative turn, that Beattie (2014) refers to, relates to combining fields such as linguistics and sociology to evaluate the role of narratives in how human actors create subjective meaning. This could, for example, relate to interpreting the whole of an annual report, that is created through text, tables, and pictures. As discussed in the article, the term narrative can be contrasted with the more objective connotation of the word disclosure, which in accounting suggests that accountants merely disclose objective facts that are not up for debate. Based on her review of relevant literature, Beattie (2014, 115) describes how documents such as annual reports have evolved from simple legal documents to "flamboyant documents exhibiting creative use of text and visual images". Such development enables or even necessitates a deeper evaluation of not just the content, but also of the context and the form of the presentation.

Research into accounting narratives could be categorized in several ways. Beattie (2014) finds five main areas. These relate to studying what is reported (descriptive studies), explaining why the observed practice is such as it is (finding the determinants), understanding the author's or narrator's explanations for the observed practice (interpretive approach), evaluating the consequences of the practices, and considering the normative implications (what should be reported). Meaningful contributions to existing knowledge on non-GAAP reporting could be achieved through each of these areas. In this thesis the focus is on the central area of what is reported, which in Beattie's (2014) categorization includes methods such content analysis, linguistic analysis, and evaluation of impression management (for example consisting of tone, deception, or persuasion).

In historical context, the use of linguistic tools such as readability measures has been one of – if not the most – used approach for applying textual analysis to accounting research. Readability measures are generally applied together with the hypothesis that the firms' disclosures are consciously made more difficult to read and understand in order to obfuscate certain negative items. Prior studies provide evidence on better readability being linked to higher valuation (Caglio et al. 2019), firm-specific information environment improving through better readability (Bai et al. 2019), and how worse readability is associated with lower accuracy and higher uncertainty of analyst earnings forecasts (Lehavy et al. 2011). The popularity of the readability measures may be explained through how easy it is to automate the measurement of these, even for large samples. Generally used readability measures focus simply on sentence length and absolute or relative number of complex words, with word complexity normally being measured simply through the number of syllables. Another recurring approach has been the evaluation of the tone of the message (e.g. Rogers et al. 2011; Guillamon-Saorin et al. 2017), where overly positive or optimistic tone can be considered indicative of management trying to inflict positive bias.

An often cited article on measuring the readability of financial disclosures is Loughran and McDonald (2014), which provides good summary on how these measures are constructed. The article also covers why these simple readability formulas are in many ways problematic when used to evaluate financial disclosures. Already previously Loughran and McDonald (2011) had discussed how simplistic analysis of tone through negative word counts is a problematic approach within finance. The gist of the authors' argument is that common lexicons or corpora, which refer to word list resources used in linguistic research, are not directly applicable within finance. Words such as vice, liability or foreign are commonly labelled negative in tone, but are neutral and naturally recurring

words in the financial lexicon. This is also a limitation with more complex and automatized natural language processing methods that rely on any non-industry-specific grammar list or corpus.

Manual content analysis can currently offer a more thorough view into textual analysis than most simplistic linguistic methods. This is also where impression management enters the picture. Content analysis as a method in narratives research is also discussed in Beattie (2014, 114–115), who described how within this method one could do the analysis both through fully or semi-quantitative analysis of prespecified items, or as a deeply holistic evaluation of entire texts and their themes. Beattie (2014, 117) also identifies a "a strong stream of research into impressions management", which in the study was understood as employing persuasion and rhetorical devices. Impression management as a term more specifically comes from psychological literature and can be understood as "the process by which people control the impressions others form of them" (Leary & Kowalski 1990, 34).

Brennan et al. (2009) have developed an accounting-specific holistic measure of impressions management that has been further validated in later studies. The method was developed especially with annual report press releases in mind. Four techniques or areas of measurement are included:

1. Thematic analysis of keywords, statements, and amounts.
2. Analysis of selectivity (of what to present).
3. Analysis of presentation (including positioning, emphasis by repetition and emphasis by reinforcement).
4. Use of performance comparisons.

As described by Brennan et al. (2009, 799), "these four techniques lend themselves to manual content analysis of disclosures and as such form a methodologically cognate cluster". Manual analysis and coding are argued for especially through the fact that impression management techniques are subtle and sophisticated: they cannot be effectively identified with simplistic, computerized methods. Brennan et al. (2009, 801) also summarize well prior studies showing how human and computer-assisted coding can lead to significantly different results in this area. This kind of analysis that incorporates several dimensions and is not reliant on too formulaic, automatized measures, makes it possible to extract more nuanced findings from the subject text, compared to something like a traditional readability measure.

The first of these techniques, thematic analysis, centers around analysis of themes or tones in the text. Brennan et al. (2009, 804) describe how this can be done at any level from individual words to paragraphs and to the whole text, though they do suggest that analysis at sentence-level can be seen as the best in terms of reliability. In applying this method, thematic analysis can be operationalized for example through analysis of positive and negative amounts, keywords and statements, as is done for example in García Osma and Guillamon-Saorin (2011) and Guillamon-Saorin et al. (2017). This kind of application typically utilizes a specific lexicon that contains a classification of positive and negative keywords.

Second on the list is analysis of selectivity. Especially outside of the main financial statements, companies can decide what information to disclose. As commented by Brennan et al. (2009, 808), in documents such as press releases firms can be assumed to present the highest possible profit or earnings figures. From a methodological point of view, researcher could for example compare the figures included in a press release to the other options available in the full set of financial statements. This kind of method has been followed by García Osma and Guillamon-Saorin (2011).

Analysis of presentation, the third technique, contains several dimensions, but can be understood especially through emphasis. The idea is that when certain information is emphasised, it becomes the main item that the reader notices and relies on for creating the initial impression. In Brennan et al. (2009) emphasis is measured through three ways: location or positioning, repetition, and reinforcement. Positioning relates to the fact that placing information in more prominent location increases the likelihood that the reader notices and pays attention to it. Emphasis through location can be measured for example by assuming that the initial headline and ingress are the most emphasised, and the last paragraph of the last page is the least emphasised (the emphasis is linearly decreasing with document length). A company might emphasise good news and include them in the heading but hide the negative facts to the last page of the document. Repetition, meaning that same piece of information appears in the document more than once, is another way for emphasis. It is self-evident that repetition makes it easier and more likely that the reader remembers and takes away that specific piece of information. Finally, Brennan et al. (2009) also include reinforcement as part of evaluation of emphasis. This means that a keyword or figure is accompanied with additional reinforcement: for example, description of growth can be reinforced by describing it as *rapid* and *profitable* growth – not just your ordinary growth. Conversely the impact of keywords or figures could be diminished

through qualifiers. A company facing falling sales could for example describe how the sales have fallen, but only *somewhat* or *little*. Examples of applying this again include studies by García Osma and Guillamon-Saorin (2011) and Guillamon-Saorin et al. (2017).

Final piece of the Brennan et al. (2009) impressions management measure is the use of performance comparisons. The idea here is that this comparative information can be used to reinforce or diminish the newer information. Company might include a benchmark or comparison only when it highlights a positive change. This could be seen in company presenting a positive percentage change when sales have decreased, but not giving any information on the change or the prior year value when sales has decreased.

Through these four techniques a comprehensive measure for impression management can be formulated. By considering more than just one simple impressions management technique and by using human coding, this measure allows for focused and detailed understanding of a company's use of impressions management. Prior studies have applied the measure and identified new and meaningful findings (García Osma & Guillamón-Saorín 2011; Guillamon-Saorin et al. 2017). The detailed application of the measure in this thesis, including how to calculate the total composite impressions management score, is discussed in chapter 4.2.

3 HYPOTHESES DEVELOPMENT

The main objectives of this thesis revolve around the quality of the non-GAAP disclosures and the adjustments made in reaching those figures. As the review of the relevant literature shows, many methods have been devised to evaluate the quality and usefulness of APMs. Possibly the most common approach has been the evaluation of the information content of the adjusted earnings, meaning that the market reaction to the figure is measured (e.g. Bhattacharya et al. 2003; Marques 2006). Here another measure of quality is used: the predictive ability of the adjusting items over future performance.

Adjusting item refers to adjustments made when deriving the alternative performance measure from the IFRS-counterpart. Specifically in Finland, FIN-FSA (2017, 10) has found that companies most commonly state that an APM is disclosed in order to present a performance measure that either is more comparable between reporting periods or gives a better understanding of the underlying, operational business performance. This means that from the regulation-based figure only the effects of transactions that are non-recurring or not related to the core operations of the firm are eliminated. Therefore, these adjustment items themselves should not have predictive ability over future performance: they are eliminated from the regulation-based figure specifically because they are stated to not reflect to normal business operations and performance. This reasoning for adjusted earnings is present elsewhere in the world as well, and the setup of modelling non-GAAP adjustments' predictive power over future performance has been used in prior studies as well (e.g. Doyle et al. 2003; Lougee & Marquardt 2004; Leung & Veenman 2018).

This approach for understanding the quality of APMs has several strengths. For one, it requires relatively little assumptions to be made, most importantly it does not require strict market efficiency assumptions. This method is also closely linked to the ESMA Guidelines, which requires that APMs must be labelled clearly. For example, a firm is not allowed to present a measure such as operating profit less non-recurring items, and then also eliminate the effect of recurring items from the original operating profit.

Based on this, the first hypothesis can be stated in the null form:

H1: *Adjustments made in deriving the non-IFRS earnings have no predictive power over future performance.*

The ESMA Guidelines came into force in July 2016. The aim of the ruleset has been to increase the transparency, neutrality and comparability of alternative performance measures (ESMA 2015). The Guidelines focus especially on presentation technical matters such as how the APMs must be clearly defined and named, also these must be reconciled to the closest IFRS line item. The APMs adopted must also be used consistently. Most importantly the rules require clear labelling of the measure, as touched on above. Relatively little guidance is given on the content of the APMs. As has been commented previously in this thesis, prior to the Guidelines there was only very limited regulation over the use of APMs in Europe.

It may be assumed that in less regulated environment the use of APMs is more aggressive. For example, the results by Abernathy et al. (2014) show that firms move from one earnings management method to another when the use of the initial method of choice becomes constricted. Isidro and Marques (2015) similarly have shown that the usefulness of APMs decreases when local corporate governance environment improves, meaning that APMs become more important earnings and impression management tool when no alternatives are available.

More direct evidence on effects of regulation on APMs exists as well. As briefly commented in chapter 1.2, in the US related guidance was introduced in the form of the Regulation G already in 2003. Several studies have evaluated the effects of the Regulation G on the use of non-GAAP figures. Initially the results suggested that the use of non-GAAP figures became less aggressive and their usefulness increased after the regulation came into effect (Jennings & Marques 2011). Later studies have found this initial effect as well, though also suggesting that the effect has since diminished (Black et al. 2012; Black & Christensen 2018). In the US, the regulatory intervention then had at least a short-to-medium term effect on the opportunistic use of APMs.

The second hypothesis can be formed as follows:

H2: *The adoption of the ESMA Guidelines increases the quality of the non-IFRS earnings measures.*

This thesis is interested in the financial communications, precisely the impression management, employed by the companies in relation to the non-IFRS measures. The prediction is that firms use impression management to manipulate the recipients' view of the true performance. This can happen through methods such as emphasizing good news and

the figures that give the most flattering view of the performance, while hiding any negative items. This can be assumed to be the case especially in the financial statement bulletins, of which content is more freeform and leaves room for management discretion.

Impression management is predicted to be more prevalent when the quality of the adjusted earnings is lower because this is a method to further persuade the investors to accept the message as it is presented. It is assumed that firms follow this kind of strategic and uniform messaging strategy, where the use of APM is one part, but also the overall messaging is adjusted to highlight and emphasize the APM. This is assumed to be the case especially when presenting low quality APMs, since it can be assumed that in such cases the firm is more reliant and aggressive with its disclosures and trying to inflict positive bias on the recipient. Guillamon-Saorin et al. (2017) have also previously shown that when excluding recurring items from non-GAAP earnings, managers are more often also engaging in high levels of impression management.

The third and final hypothesis can be formed as follows:

H3: *When accompanied with higher levels of impression management, the non-IFRS earnings measures are of lower quality.*

4 METHODOLOGY AND DATA

4.1 Sample selection and gathering the data

This thesis utilizes annual earnings press releases (also referred to as financial statements bulletins, reviews, and releases) as the base for the empirical analysis. For companies listed on the Nasdaq Helsinki stock exchange, this is a mandatory report and centrally available from the local centralized storage facility (also known as officially appointed mechanism, OAM). The initial sample consists of all such reports included in the OAM for the companies listed on Nasdaq Helsinki for the years 2012 to 2018. The documents are downloaded and reviewed to identify any possible missing observations. Missing observations are again searched from the system, and all but four of the missing firm-year observations are found. The gathered sample is compared to another listing of the constituents of the stock exchange (retrieved from the Datastream database) to identify any missing companies. No companies are found to be missing from the dataset gathered from the OAM.

Thus, the initial sample contains 876 firm-year observations for 150 different companies. This means approximately 120 observations per year, with only limited yearly variation. From this initial sample, first all financial companies are excluded. Financial companies are subject to differing regulation and typically use different performance metrics than companies in other industries. The disclosures are therefore not considered to be comparable enough with the wider sample. This reduces the sample size by 94 firm-year observations. Next all observations with only Finnish disclosures are eliminated from the sample. This is done to allow uniform application of the impression management score. While the application of this score is not strictly language dependent and it could be used with Finnish disclosures as well, for comparability of the coding only one language needs to be used. English is chosen due to the fact that this is the language in which the score has been developed and tested previously. This choice also allows better comparability to prior studies. Most listed Finnish companies issue their financial releases in English. When eliminating the observations with only Finnish documents, the sample is reduced by 72 observations, with the total remaining at 683.

For remaining companies in the sample, financial information is gathered from Eikon and Datastream databases. This includes information on the companies' size, revenue, profitability and more. Variables included in the empirical tests are discussed in chapter

4.3. The databases are missing information for 31 companies. Due to missing datapoints, the sample is further reduced by 97 firm-year observations. Table 2 portrays at more detail how the different methodological choices affect the sample. After these choices, the sample remains at 586 observations, which distribute fairly evenly between the different years. The sample contains 321 observations pre-2016 (the adoption of the ESMA Guidelines) and 265 observations post-2016. At this point 60% of the original observations remain.

Table 2 Sample size reduction due to methodological choices

	TOTAL	2012	2013	2014	2015	2016	2017	2018
INITIAL SAMPLE	**849**	**117**	**118**	**117**	**120**	**124**	**125**	**128**
LESS FINANCIAL COMPANIES	-94	-11	-11	-12	-14	-15	-15	-16
LESS ONLY NON-ENGLISH DOCUMENTS	-72	-12	-9	-10	-8	-10	-11	-12
LESS COMPANIES WITH MISSING DATA	-97	-16	-17	-16	-15	-14	-10	-9
LESS DOCUMENTS NOT CONTAINING AN APM	-233	-35	-34	-31	-30	-33	-34	-36
FINAL SAMPLE	**353**	**43**	**47**	**48**	**53**	**52**	**55**	**55**

The remaining 586 sample documents are reviewed to identify reported non-IFRS earnings measures. Manual collection of the reported APMs directly from the reporting firms' disclosures is the best option for the validity of the results. Historically research on non-GAAP reporting has often used other proxies for the non-GAAP figures disclosed by the firms: most common option is to use the I/B/E/S analyst version of the figure. As summarized by Black et al. (2018, 274–275), this approach has been criticized for not being a very good proxy for the firms' actual disclosures. Most importantly, Bentley et al. (2018) found there to be especially notable difference in the quality of the adjustments made by the analysts versus the management, with the latter having the lower quality. Collecting the data directly from the reports issued by the firms can therefore be seen vital when studying the quality of the firms' non-IFRS metrics.

Only adjusted EBIT or operating profit figures are collected for the tests: other APMs are not included in the scope of this study. It is important to note that EBIT or operating profit is in by itself an APM as per the ESMA Guidelines, even when calculated as a subtotal of IFRS line items. This is because IFRS does not include any definition of operating profit or EBIT. However, for the purposes of this thesis, only adjusted versions of these metrics are included in the scope. This is because the focus of the evaluation is on

the discretionary adjustments made to the regulation-based figures, not in simple IFRS-based subtotals.

Only the APMs included in the downloaded annual releases are considered. All the documents are read through to identify mentions or use of adjusted earnings measures. In case of unclear figures, the basis of calculation for the figure is reviewed. The APMs are typically included in the beginning of the documents and presented both together with the key figures and with the review of annual performance. Manual review is important because the naming of APMs varies significantly between companies, and simply searching for certain names may not identify all disclosed figures. Nevertheless, the documents are also searched for any mentions of common APM names. This includes going through all mentions of adjusted, comparable, operating, operational, alternative (which are common prefixes of APMs); and going through all mentions of EBIT in the document. Through these actions, only very limited possibility exists that an adjusted earnings measure would be missed in reviewing a document.

This thesis evaluates the non-IFRS reporting practices, due to which all observations without an APM are eliminated from the sample for the empirical tests. After this step, the final sample contains 353 firm-year observations, ranging from 43 to 58 per year. The number of APMs reported is seen to increase steadily during the sample period. APMs were searched from documents of 98 unique companies, of which 76 report at least one APM during the sample period. On the level of individual firm-year observations, 60% are seen to contain adjusted operating profit metric.

Next the coding of the impression management score is detailed for the documents remaining in the final sample.

4.2 Constructing the impression management score

The applied impressions management score directly follows that implemented by Guillamon-Saorin et al. (2017), who similarly used the score in relation to non-GAAP reporting. The theoretical background of the score is developed by Brennan et al. (2009) and the score has also been applied by Garcia Osma and Guillamon-Saorin (2011). The coding is based on manual content analysis: as discussed in chapter 2.5, the impressions management techniques can be very subtle, and review of items such as performance comparisons is difficult to automate.

The applied score is based on three distinct impressions management techniques, which are measured through several distinct sub-measures. The three techniques are tone,

emphasis, and performance comparisons. The setup is summarized in Table 3. From the original model (discussed in chapter 2.5) selectivity is left out of the evaluation, because the studied phenomenon (APMs) is not based on the financial statements. This means that the figure is not something that can be picked for presentation from the normal dataset. The coding is done directly in relation to the identified APMs and observations without any APMs are not included in the scope. This approach is aligned with Guillamon-Saorin et al. (2017). The coding is done on the most and next-most emphasized sections of the documents, also aligned with Guillamon-Saorin et al. (2017). This focuses the attention to the most emphasized and important parts of the document, since this is what recipients are considered to use for forming their opinion and what grasps the attention of the news outlets (Entwistle et al. 2006; Guillamon-Saorin et al. 2017).

Table 3 Methods of measuring impression management, adapted from Guillamon-Saorin et al. (2017, 457)

TECHNIQUE	OBJECT OF TECHNIQUE	MEASURE
TONE	Keywords	Number of positive and negative keywords
	Quantitative amounts	Number of quantitative positive and negative amounts
EMPHASIS	a) Location/positioning/presentation of keywords	Most and next-most emphasized section
	a) Location/positioning/presentation of amounts	Most and next-most emphasized section
	b) Repetition of statements	Number of positive and negative repetitions of statements
	b) Repetition of quantitative amounts	Number of positive and negative repetitions of amounts
	c) Reinforcement of keywords	Number of positive and negative repetitions of reinforcements
PERFORMANCE COMPARISONS	Quantitative amounts	Benchmark, previous year amount comparison, both

Tone refers to the amount of positivity in the communications - as opposed to merely neutral disclosures. The analysis of tone is based on analysis of keywords and quantitative amounts related to the disclosed non-IFRS figures. Keywords used in relation to an APM are labelled positive or negative, for which the word sentiment list by Loughran and McDonald (2011) is used as a background source. This finance and accounting focused word sentiment list has been updated after the initial release in 2011 and the 2018 version is used here. The latest word list is publicly available from the University of Notre Dame's website. As an example of an instance leading to coding of a positive keyword, Glaston Oyj Abp's document from 2018 starts with the headline "Good fourth quarter - orders

received grew and comparable operating profit improved": here the word "improved" is coded as a positive keyword related to an adjusted earnings figure. The quantitative non-IFRS amounts are classified as positive (negative) if they exceed (are lower than) the comparison year's figure. Alternatively, if the document explicitly states the figure to be positive or negative, then this classification is used.

Emphasis is about making the specific piece of information more noticeable and obvious to the recipient. Three measures are used for identifying emphasis. The first measure is the positioning of the keywords or quantitative amounts. This is emphasis by location, and it can be understood such that strategically presenting something earlier in the document enhances the visibility of that piece of information – after all, readers are first paying attention to the headline and following summary. Above example from Glaston Oyj Abp's document from 2018 included a positive mention of an APM in the headline – meaning a mention in the most emphasized section of the document. The second measure is repetition of statements or amounts, meaning that specific APM or related positive or negative statements are repeated in the document. For example, Finnair Oyj's document from 2017 mentions comparable operating result of 170 million euros first in the headline and then again both in the annual summary and in the CEO comment. While it is common that information included in headline is repeated in the main text, this is still considered emphasis by repetition. The company has after all first decided to include the headline, then to include the specific figure in that headline, and then to repeat the figure also in the main text. The third measure of emphasis is reinforcement of keywords, which means including qualifiers to emphasize the connotation. In Exel Composites Plc's financial statements release for 2017 already the headline states that "revenue and operating profit increased significantly", where the word "significantly" is positive reinforcement of *how* the operating profit has increased.

Performance comparisons refer to the firms' decision to include benchmark information for the APMs. This does not cover presenting the comparison period's figure in brackets behind the current year's figure, but rather additional benchmarks. For example, in 2013, Fiskars decided to mention that "operating profit excluding non-recurring items increased 17% to EUR 73.8 million", where the 17% is coded as a positive performance comparison.

The composite impression management score is calculated from the above measures and the calculation logic is summarized in Table 4. Each keyword and amount are given a weight of 1. If a keyword or amount appears in the most or next-most emphasized

section, a weight of 1 or 0.5, respectively, is added. For reinforced keywords, a weight of 0.5 is added. Similarly, a single instance of performance comparison or repetition adds a weight of 0.5. The weights are negative or positive depending on whether the item is coded as negative or positive. The total composite impression management score is the sum of the positive and negative weights divided by the number of words in the section. The scaling is included to ensure the comparability of disclosures of different length.

Table 4 Calculation of the composite impression management score

TECHNIQUE	OBJECT OF TECHNIQUE	WEIGHT
TONE	Keywords	+1 for positive, -1 for negative
	Quantitative amounts	+1 for positive, -1 for negative
EMPHASIS	A) Location/positioning/ Presentation of keywords	0.5 / -0.5
	A) Location/positioning/ Presentation of amounts	0.5 / -0.5
	B) Repetition of statements	0.5 / -0.5
	B) Repetition of quantitative amounts	0.5 / -0.5
	C) Reinforcement of keywords	0.5 / -0.5
PERFORMANCE COMPARISONS	Quantitative amounts	0.5 / -0.5
TOTAL SCORE		**SUM OF WEIGHTS DIVIDED BY TOTAL NUMBER OF WORDS IN SECTIONS CODED.**

4.3 Research design

The hypotheses from chapter 3 focus on the quality of the adjustments made in deriving the non-IFRS figure. The quality of the adjustments is measured with their persistence. As has been described previously in the thesis, the repeating justification for the adjusted earnings figures is that these present the core, recurring result of the firm. Higher persistence of the earnings adjustments is hypothesized to be linked to lower quality adjustments, just as in previous studies (e.g. Doyle et al. 2003; Frankel et al. 2011; Guillamon-Saorin et al. 2017).

Following model is estimated, similar ones used by Guillamon-Saorin et al. (2017) and Frankel et al. (2011):

$$Future\_Performance_{i,t+1} = \alpha_0 + \beta_1 ADJ\_EBIT_{i,t} + \beta_2 Adjustment_{i,t} + \beta_3 Post\_ESMA_{i,t} + \beta_4 High\_IM_{i,t} + \beta_5 Post\_ESMA_{i,t} \times Adjustment_{i,t} + \beta_6 High\_IM_{i,t} \times Adjustment_{i,t} + \beta_7 Size_{i,t} + \beta_8 Loss_{i,t} + \beta_9 Earnings\_Vol_{i,t} + TimeControls + u_{i,t}.$$

The dependent variable $Future\_Performance$ for firm *i* in period *t+1* is operationalized both through *operating result* and *net cash flow from operations*. Estimating the model separately both for earnings and cash flows ensures a more comprehensive view on the persistence of non-IFRS adjustments. This approach has been previously recommended by Whipple (2015) and later by Black et al. (2018), and applied by Leung and Veenman (2018).

In a more traditional test settings, the main independent variable could be the IFRS earnings, in which case the results would indicate whether current period's IFRS earnings have predictive power over future earnings. Here the focus is on non-IFRS earnings, and thus the first independent variable is ADJ_EBIT, which is the non-IFRS operating result reported in the firm *i*'s annual earnings release in year *t*. The second independent variable of interest is $Adjustment$, which is the IFRS operating result minus the non-IFRS version, in essence the adjustment made to the IFRS-compliant figure. If the non-IFRS measure presents the result of recurring operations, then the non-IFRS adjustments should not have predictive power over future performance, meaning the coefficient for $Adjustment$ would not statistically differ from zero. Based on previous results of Guillamon-Saorin et al. (2017) and Frankel et al. (2011), it is assumed that the adjustments contain also recurring items, and the coefficient does differ from zero. $Future\_Performance$, ADJ_EBIT, and $Adjustment$ are measured in millions of euros, allowing for easy interpretation of coefficients for these variables and related interactions.

The main remaining variables of interest are the interactions $Post\_ESMA \times Adjustment$ and $High\_IM \times Adjustment$. $Post\_ESMA$ is a dummy variable, with value of 1 for observations post-2016, after the ESMA Guidelines came into effect, and 0 otherwise. The interaction $Post\_ESMA \times Adjustment$ indicates whether the ESMA Guidelines have had any effect on the quality of the non-IFRS adjustments. It is expected that the ESMA Guidelines would have improved the quality, as happened in the US after corresponding regulation was introduced (Jennings & Marques 2011). In such case the coefficient for the interaction term should be closer to zero than the one for $Adjustment$. $High\_IM$ represents the impression management score such that the variable gets value of 1 when the impression management score for the observation is higher than the sample median score, and 0 otherwise. The interaction $High\_IM \times Adjustment$ represents the persistence of non-IFRS adjustments when the APM is accompanied with high levels of impression management. The expectation is that the coefficient is significantly negative,

aligned with the hypothesis that impression management is used to hide the low-quality adjustments when firms try to obfuscate the true performance. This is aligned with the results by Guillamon-Saorin et al. (2017).

$Size$, $Loss$, and $Earnings\_Vol$ are control variables similar to ones used in prior research (see e.g. Doyle et al. 2003; Frankel et al. 2011; Guillamon-Saorin et al. 2017). $Size$ is the natural logarithm of the total assets and controls for the effect of size. $Loss$ is a dummy-variable, having value of 1 if the net profit is negative and 0 otherwise. The persistence of earnings may differ between profit and loss-making years. $Earnings\_Vol$ is measured as ROA volatility over the past 3 years, and controls for recurring volatility in profitability. Due to nature of panel data, the model is estimated with time fixed effects, represented in the model by $TimeControls$. The fixed effect for time controls for firm-specific time-invariant characteristics.

The OLS regression model is estimated with robust standard errors as described by Hinkley (1977) to combat issues of heteroscedasticity. The Hinkley (1977) robust standard errors improve on the normal robust standard errors by improved small sample properties, though the methods are asymptotically equivalent (Hayes & Cai 2007). Long and Ervin (2000) have provided evidence that the small sample properties of the robust standard errors are especially relevant for sample sizes below 250, but are less relevant above that, as the asymptotic properties enter the picture. With sample size of over 300, using heteroscedasticity corrected standard errors can anyhow be considered warranted. While homoscedasticity can be considered the most important assumption of OLS for statistical interference, the OLS regression also assumes normal distribution of the error terms. Through central limit theorem the OLS estimators lead to the underlying errors approaching normality as the sample size increases. No absolute rule exists for when asymptotic normality can be assumed, but some suggestions include sample size of over 30 or considering the degrees of freedom through an equation of sample size less independent variables less one. (Wooldridge 2013, 175–177) With sample size of over 300, assumption of asymptotic normality of error terms is applied.

5 EMPIRICAL RESULTS

5.1 Descriptive statistics and univariate tests

Table 5 presents the descriptive statistics for all the variables included in the model. The main variables are measured in millions of euros. Of the main variables, it can be seen that the range for operating result and cash flows is wide. The notable difference between sample mean and median for these variables means that certain outlier results are inflating the mean away from the midpoint of the distribution. This can be explained by the nature of the companies listed on Nasdaq Helsinki, with there being notable size-variation between the index constituents. For example, in 2019 for half of the companies the revenue totaled 268 million euros or less, but the highest revenue was 23,315 million euros. Firm size is controlled in the model with the control variable $Size$, which is the natural logarithm of total assets. Since the model is estimated with time fixed effects, firm-specific time-invariant characteristics are also controlled for.

Table 5 Descriptive statistics

	MEAN	1ST QUARTILE	MEDIAN	3RD QUARTILE	STANDARD DEVIATION
FUTURE_PERFORMANCE (EBIT)	149.299	2.311	24.000	140.000	389.083
FUTURE_PERFORMANCE (CF)	187.969	5.465	41.920	180.900	384.497
ADJ_EBIT	179.444	4.600	31.300	164.200	371.615
ADJUSTMENT	38.131	0.057	3.075	20.100	279.173
POST_ESMA	0.459	0.000	0.000	1.000	0.499
HIGH_IM	0.499	0.000	0.000	1.000	0.501
POST_ESMA X ADJUSTMENT	18.382	0.000	0.000	1.591	200.760
HIGH_IM X ADJUSTMENT	23.920	0.000	0.000	2.900	187.297
SIZE	6.322	4.605	6.292	7.850	1.962
LOSS	0.262	0.000	0.000	1.000	0.440
EARNINGS_VOL	1,950.495	10.304	43.655	203.272	17,385.045

The mean for non-IFRS operating result is 179.4 million euros, while the median is 31.3 million euros. The first quartile value for the corresponding variable ADJ_EBIT is 4.6 million euros. In 41 cases ADJ_EBIT is negative, while the normal IFRS-based version is negative in 71 cases. Approximately one fifth of the sample thus represents negative result from operations. The relatively low number of observations with operating losses

could be explained by the nature of listed companies, which typically are more mature companies, for which straight out operating losses are not as common.

The mean for $Adjustment$ is 38.1 million euros, while the median is 3.1 million euros. The first quartile value of 0.1 million euros is still positive. In a clear majority of cases the adjusted profit is higher than the IFRS counterpart. This is consistent with prior evidence showing that the adjustments are mostly concerning expenses, as discussed in chapter 2. For half of the observations the non-IFRS measure exceeds the IFRS version by over 3 million euros. Through the interaction term $Post\_ESMA \times Adjustment$ it can be seen that in the post-ESMA Guidelines era the average adjustment is lower than pre-2016. The average adjustment is also lower when there are high levels of impressions management in place, as seen from the interaction term $High\_IM \times Adjustment$, which represents the non-IFRS adjustment in cases where high levels of impression management are in place. The standard deviations are fairly big due to the somewhat long-tailed distribution.

During the sample period the number of companies reporting a non-IFRS operating result figure increased by over 30%, which does not account for possible yearly fluctuations. Of the 98 companies included in the analysis, 21 (or 21%) did not report APMs at all during the sample period. On the other hand, 42 companies (43%) included an APM in all of their reports in the sample period. The reporting frequency of non-IFRS metrics is seen to be increasing over time, reflecting the previously identified trend.

Table 6 presents the Pearson correlations for all the variables included in the model. The correlation between the non-IFRS figure and following year's performance is positive and significant at 0.553 and 0.663 for future EBIT and cash flow, respectively. This is aligned with the expectation that the core performance is to some extent fixed over the years. On the other hand, the correlation between the non-IFRS adjustment and future performance is not statistically different from zero, though the correlation coefficient is negative when future performance is measured by EBIT. The negative correlation implies that the non-IFRS exclusions actually predict future expenses, implying a lower quality disclosure. Statistically significant and negative correlation is found for both $Post\_ESMA \times Adjustment$ and $High\_IM \times Adjustment$ against future performance when measured with EBIT, with respective values of -0.177 and -0.133. This indicates that higher levels of impression management are linked to lower quality disclosures, as is expected. Curiously also the post-ESMA Guidelines periods are seen to be linked to lower quality disclosures, indicating that after the implementation of the regulation the reporting

quality has gotten worse. When future performance is measured with cash flows, the correlation coefficient for these terms is negative, but not statistically different from zero.

Table 6 The Pearson correlations

Asterisk (*) indicates statistical significance at the .05 significance level. The variables are defined in chapter 4.3.

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)
(1) *FUTURE_ PERFORMANCE(EBIT)*	1.000										
(2) *FUTURE_ PERFORMANCE(CF)*	0.744*	1.000									
(3) *ADJ_EBIT*	0.553*	0.663	1.000								
(4) *ADJUSTMENT*	-0.073	0.081	0.368*	1.000							
(5) *POST_ESMA*	-0.039	0.017	0.064	0.007	1.000						
(6) *HIGH_IM*	0.013	0.013	0.110*	0.036	0.078	1.000					
(7) *POST_ESMA X ADJUSTMENT*	-0.117*	-0.003	0.424*	0.713*	0.101*	0.025	1.000				
(8) *HIGH_IM X ADJUSTMENT*	-0.133*	-0.078	0.465*	0.664*	0.007	0.130*	0.650*	1.000			
(9) *SIZE*	0.491*	0.601*	0.640*	0.199*	0.022	0.122*	0.155*	0.202*	1.000		
(10) *LOSS*	-0.215*	-0.163*	-0.164*	0.175*	-0.117*	-0.205*	0.136*	0.089*	-0.336*	1.000	
(11) *EARNINGS_VOL*	-0.043	-0.055	-0.050	-0.011	0.044	0.013	-0.007	-0.008	-0.160*	0.166*	1.000

Between independent variables the highest correlations are between ADJ_EBIT and $Size$ (0.640) along with ADJ_EBIT and the two main interaction terms $Post\_ESMA \times Adjustment$ and $High\_IM \times Adjustment$ (0.424 and 0.465). Inherently the interaction terms are strongly correlated with the underlying covariates. Variance inflation factors (VIF) of the independent variables are presented separately in Table 7. No excessive collinearity is identified based on the VIF values either, with the highest value remaining below three. OLS assumption of no perfect collinearity between independent variables is not considered to be violated.

Table 7 Collinearity statistics

	VIF
ADJ_EBIT	2.237
ADJUSTMENT	2.466
POST_ESMA	1.048
HIGH_IM	1.080
POST_ESMA X ADJUSTMENT	2.462
HIGH_IM X ADJUSTMENT	2.232
SIZE	1.971
LOSS	1.294
EARNINGS_VOL	1.054

5.2 Results for the main tests

The results from the main regression model are presented in Table 8. As described in chapter 4.3, the main goal is to test whether the non-IFRS adjustments have predictive power over future performance, with future performance operationalized both through future operating profit and future operating cash flows. Results for both operationalizations are presented in Table 8.

Hypothesis 1 states that adjustments made in deriving the non-IFRS earnings measure have no predictive power over future performance. This would be aligned with the comments of the reporting firms, who most commonly state that the non-IFRS earnings figure seeks to describe the core or recurring result of the firm. Several prior studies have found non-GAAP earnings adjustments to predict future performance (e.g. Doyle et al. 2003; Frankel et al. 2011; Guillamon-Saorin et al. 2017), but as seen through chapter 2, the results on aggressive or opportunistic use of non-GAAP figures overall is more mixed.

For dependent variable $Future\_Performance_{EBIT}$ the estimated coefficient for $Adjustment$ is 0.070 but the coefficient is not statistically different from zero with p-value of 0.805. For dependent variable $Future\_Performance_{CF}$ the coefficient is higher at 0.293, though with p-value of 0.111 it is not statistically different from zero. Results for both dependent variables support each other. The coefficients for independent variable ADJ_EBIT are statistically significant at the 0.01 level for both operationalizations of future performance. In essence one euro of non-IFRS earnings predicts 0.773 euros of operating earnings and 0.806 euros of operating cash flow for the following year. The non-IFRS adjustments have no bearing on future performance. The results support accepting hypothesis 1. This also supports the reporting firms' arguments that non-IFRS earnings present the core result of the business and that only one-time items are excluded from the discretionary measure. This result is opposite to what some prior studies from the US and wider European setting have shown. For example, Doyle et al. (2003) found that the non-GAAP earnings adjustments predicted future cash flows at a statistically significant level in the US. Guillamon-Saorin et al. (2017) similarly found non-IFRS earnings adjustments to predict future earnings in Europe during their sample period from 2003 to 2009.

Table 8 Results of the main regression model

The OLS regression model is estimated with robust standard erros as described by Hinkley (1977). The model includes time fixed effects based on calendar years to control for firm-specifc time-invariant characteristics. All P-values are two-sided. Asterisk (*) indicates statistical significance at the .05 significance level and double-asterisk (**) at the .01 significance level. The variables are defined in chapter 4.3.

Panel A
Dependent variable $Future\_Performance_{EBIT}$

Variable	Coefficient	Robust standard error	t-value	P-value
INTERCEPT	-111.170	76.754	-1.448	0.148
ADJ_EBIT	0.773	0.107	7.231	0.000**
ADJUSTMENT	0.070	0.282	0.247	0.805
POST_ESMA	-70.841	72.000	-0.984	0.326
HIGH_IM	-21.146	33.005	-0.641	0.522
POST_ESMA X ADJUSTMENT	-0.433	0.272	-1.589	0.113
HIGH_IM X ADJUSTMENT	-0.800	0.242	-3.305	0.001**
SIZE	25.005	10.156	2.462	0.014**
LOSS	-13.720	27.439	-0.500	0.617
EARNINGS_VOL	0.000	0.000	2.029	0.043*
R2	0.550			
Adjusted R2	0.531			

Panel B
Dependent variable $Future\_Performance_{CF}$

Variable	Coefficient	Robust standard error	t-value	P-value
INTERCEPT	-245.886	64.214	-3.829	0.000**
ADJ_EBIT	0.806	0.101	7.950	0.000**
ADJUSTMENT	0.293	0.183	1.598	0.111
POST_ESMA	8.468	37.976	0.223	0.824
HIGH_IM	-15.679	21.898	-0.716	0.474
POST_ESMA X ADJUSTMENT	-0.390	0.399	-0.976	0.330
HIGH_IM X ADJUSTMENT	-1.012	0.373	-2.715	0.007**
SIZE	42.170	9.563	4.410	0.000**
LOSS	50.041	25.918	1.931	0.054
EARNINGS_VOL	0.000	0.000	1.372	0.171
R2	0.699			
Adjusted R2	0.686			

Further hypotheses relate to quality of non-IFRS adjustments in specific conditions. Hypothesis 2 expects that the adoption of the ESMA Guidelines has increased the quality of the non-IFRS measures. Prior studies from the US have shown evidence that regulation may have a positive effect on the quality of these adjusted figures – at least temporarily after the implementation (Jennings & Marques 2011; D. E. Black et al. 2012). Here the independent variable $Post\_ESMA$ is a dummy-variable that equals zero when the observation is from a reporting period after the implementation of the ESMA Guidelines. The coefficient for the interaction term $Post\_ESMA \times Adjustment$ is -0.433 for operating earnings and -0.390 for operating cash flow, but in both cases the coefficient fails the test of statistical significance, both results thus supporting each other. Because the coefficient does not statistically differ from zero, the results suggest that the ESMA Guidelines have not had any effect on the quality of non-IFRS earnings reporting, since the individual coefficient for $Adjustment$ does not statistically differ from zero either. Hypothesis 2 is not supported by the results.

Finally, hypothesis 3 posits that higher levels of impressions management are linked to lower quality non-IFRS earnings adjustments. For dependent variable $Future\_Performance_{EBIT}$ the coefficient for $High\_IM \times Adjustment$ is -0.800 with a p-value of 0.001. For dependent variable $Future\_Performance_{CF}$ the coefficient for $High\_IM \times Adjustment$ is -1.012 with p-value of 0.007. This means that when a company engages in aggressive impression management, one euro of excluded items predicts negative 0.800 euros of future operating earnings and negative 1.012 euros of future operating cash flow. In presence of high impression management, the non-IFRS exclusions are of significantly lower quality, having significant predictive power over future performance. Hypothesis 3 is supported by the results. The findings are in line with prior findings of Guillamon-Saorin et al. (2017).

For both versions of the estimated model the R-squared and the adjusted R-squared measures are close to each other. When predicting future earnings, the adjusted R-squared is 0.531, while for future cash flows version it is 0.686. These values are in the same range as previous studies employing very similar models, such as the ones by Frankel et al. (2011) and Guillamon-Saorin et al. (2017). The adjusted R-squared values suggest that the model is able to explain over half of the variability of the dependent variable.

5.3 Additional analysis

The main regression model is also estimated with the quantile regression method, often credited to Koenker and Bassett (1978). While OLS models the conditional mean of the sample and utilizes squared error terms, quantile regression estimates the conditional median and uses absolute errors. Quantile regression can also be applied at other points of the distribution, including points other than the median. This makes it possible to identify differences in the coefficients in different quantiles of the dependent variable. Equivalently to traditional comparison of mean and median, quantile regression can be considered better in being robust to outliers that may distort the usefulness of an average as a proxy for the sample. (Koenker & Hallock 2001; Beyerlein 2014; Distante et al. 2018)

One benefit of the method is in identifying possible deviation in the coefficients at different quantiles of the sample, without having to split the sample to many subsamples and losing degrees of freedom. For the tests at hand, the predictive power of non-IFRS adjustments, ESMA regulation and impression management could vary at different quantiles of future performance.

Here quantile regression is used especially as an additional method for validating the results of the main test, particularly to evaluate whether results from modelling the median and the different percentiles or quantiles support the same conclusions as modelling the mean does. Extreme values of the outcome distribution and the possible trends in behavior of the quantile regression coefficients are not directly linked to the main objectives of this thesis, and thus they are considered only a secondary concern.

Table 9 presents the results from the quantile regression model for the sample quartiles as well as the first and ninth decile, with a comparison to the OLS coefficients from the previous section. At the sample median the results for the variables of interest are similar to OLS. For dependent variable $Future\_Performance_{EBIT}$ the coefficient for $Adjustment$ is very close to zero at negative 0.026 (OLS 0.070, not statistically different from zero), while for ADJ_EBIT the coefficient is again statistically significant. The coefficients for both of the interaction terms in the model similarly support the OLS results. The results at the sample median are similar for $Future\_Performance_{CF}$, with the estimated coefficients mainly supporting the OLS results. Only significant deviation is identified in the interaction term $Post\_ESMA \times Adjustment$, where the coefficient is statistically significant and clearly more negative than the OLS result, where the coefficient was not statistically different from zero. The negative coefficient suggests that the ESMA

Guidelines would have decreased the quality of the non-IFRS disclosures, which is the opposite of the expectation of hypothesis 2.

Table 9 The results of the quantile regression model

Goodness of fit for the OLS results is normal R-squared and for the quantile regression pseudo R-squared, since normal R-squared cannot be calculated. All P-values are two-sided. Asterisk (*) indicates statistical significance at the .05 significance level and double-asterisk (**) at the .01 significance level. The variables are defined in chapter 4.3.

Panel A
Dependent variable $Future\_Performance_{EBIT}$

Variable	OLS	q=0.1	q=0.25	q=0.5	q=0.75	q=0.9
INTERCEPT	-111.170	13.592	-14.399*	-21.416**	-12.183	-26.792
ADJ_EBIT	0.773**	0.407**	0.563**	0.730**	1.062**	1.507**
ADJUSTMENT	0.070	-0.189	-0.272**	-0.026**	0.130**	0.130**
POST_ESMA	-70.841	-1.259	-2.857	-3.072	-0.210	-5.809
HIGH_IM	-21.146	-2.165	-1.573	-0.605	0.962	1.175
POST_ESMA X ADJUSTMENT	-0.433	0.013	-0.078**	-0.492**	-0.945**	-0.467**
HIGH_IM X ADJUSTMENT	-0.800**	-0.550**	-0.538**	-0.529**	-0.531**	0.038
SIZE	25.005**	-3.982**	3.653**	5.612**	3.306**	7.942*
LOSS	-13.720	-5.466	-1.697	0.490	4.299	10.654
EARNINGS_VOL	0.000	0.000	0.000	0.000	0.000	0.000
Goodness of fit	0.550	0.221	0.328	0.463	0.581	0.647

(goodness of fit for OLS is R-squared, for quantile regressions the measure is pseudo R-squared)

Panel B
Dependent variable $Future\_Performance_{CF}$

Variable	OLS	q=0.1	q=0.25	q=0.5	q=0.75	q=0.9
INTERCEPT	-245.886**	6.828	-9.133	-36.415**	-74.169**	-120.634**
ADJ_EBIT	0.806**	0.404**	0.764**	1.005**	1.086**	1.332**
ADJUSTMENT	0.293	-0.038**	0.187**	0.350**	0.654**	0.614**
POST_ESMA	8.468	1.061	1.904	-0.128	-2.342	-11.198
HIGH_IM	-15.679	-4.061	-4.399	-7.291	-6.696	14.324
POST_ESMA X ADJUSTMENT	-0.390	0.216**	-0.366**	-0.784**	-1.197**	-1.443**
HIGH_IM X ADJUSTMENT	-1.012**	-1.382**	-0.643**	-0.626**	-0.515**	-0.601**
SIZE	42.170**	-1.742*	2.095	9.324**	19.526**	33.465**
LOSS	50.041	-1.822	2.321	7.599	15.179	36.515
EARNINGS_VOL	0.000	0.000	0.000	0.000	0.000	0.000
Goodness of fit	0.699	0.247	0.357	0.556	0.684	0.758

(goodness of fit for OLS is R-squared, for quantile regressions the measure is pseudo R-squared)

The characteristics of the coefficients at different points of the distribution provide some additional insights. First notion is that, aligned with above, it can be seen that at above the distribution median the coefficient for $Post\_ESMA \times Adjustment$ turns negative at a statistically significant level. Second point relates to coefficient for $Adjustment$, which is significantly negative at the lowest quartile against dependent variable $Future\_Performance_{EBIT}$. The coefficient for $High\_IM \times Adjustment$ is mostly very stable across the distribution in both of the cases.

The results of the quantile regression overall support those of the OLS regression, especially when comparing the OLS results to the sample median results. Also with the quantile regression, the results for both of the dependent variables overall agree with each other, increasing the validity of the results.

6 CONCLUSIONS

6.1 Discussion of the results

The empirical tests examined the quality of non-IFRS earnings adjustments, both overall in the sample and in relation to related regulation and the presence of impression management. The descriptive statistics highlight the increasing popularity of non-IFRS reporting in the Finnish sample, with the number of companies reporting a non-IFRS operating result figure increasing over 30% during the sample period. This is aligned with previous results on the increasing use of these adjusted figures. Studying the non-IFRS adjustments more closely, the mean adjustments is seen to be over 38.1 million euros, while the median adjustment is 3.1 million euros. This suggests that typically the adjustments are several million euros, but also some big outliers exist, dragging the average high above the median value. The difference between mean and median can be at least partly explained by the size differences between the sample companies.

Three hypotheses were formed in chapter 3, with hypothesis 1 relating to persistence of the non-IFRS earnings adjustments. Since companies use non-IFRS measures to present earnings figures that are more comparable between reporting periods or that are more presentative of the core result of the business (e.g. FIN-FSA 2017, 10), the non-IFRS adjustments should cover only non-recurring items. The adjustments should have no bearing over future performance. As seen in chapter 2, the results on benefits versus potential malpractice are indecisive, but several studies have shown non-GAAP adjustments having predictive power over future performance (e.g. Doyle et al. 2003; Lougee & Marquardt 2004; Leung & Veenman 2018).

Here the empirical results supported the reporting companies' arguments: the non-IFRS adjustments among the Finnish sample did not have any statistically significant predictive power over future performance. As has been previously recommended (Whipple 2015; Black et al. 2018), in the tests future performance was operationalized both through future earnings and future cash flows. With both measurements of future performance, the OLS coefficients for non-IFRS adjustments were not statistically different from zero. Modelling the conditional median with quantile regression also overall supported the same conclusions.

The null hypothesis of non-IFRS earnings adjustments not predicting future performance is therefore supported by the results. This is in contrast to several previous studies

suggesting the opposite. For example, Doyle et al. (2003) found that in their sample of US companies, one dollar of non-GAAP earnings exclusions predicted over 7 dollars of future cash from operations over the following three years. Frankel et al. (2011) on the other hand found one dollar of quarterly non-GAAP exclusions to predict over 1.30 dollars of expenses over the following four quarters. The samples of these studies are mostly from before the use of non-GAAP figures was regulated in the US. Guillamon-Saorin et al. (2017) used a European sample from 2003 to 2009 and found one euro of non-GAAP exclusions to predict 0.703 euros of future expenses.

Several explanations can be considered to explain differences to previous results. As seen through chapter 2, plenty of evidence exists also in favor of the usefulness of non-IFRS reporting and for the figures offering valuable information. The existing literature does therefore leave room also for results in support of good reporting quality, which is one explanation for the empirical results of this thesis. Alternatively, the sample from the past 10 years is newer than most of the preceding studies. Non-GAAP reporting has become more common since the late 1990s and one conceivable explanation is also that the quality of non-GAAP reporting has increased as users of financial information have become more aware and educated on how these figures might be misused. Testing this explanation is one potential future research topic.

Prior studies have also predominantly focused on the US. Guillamon-Saorin et al. (2017) did study the wider European setting, though over half of the sample consisted only of companies from the UK and France. The difference to previous results could therefore also be explained by country differences since the Finnish setting specifically has not been studied previously. Isidro and Marques (2015) have provided evidence that aggressive or opportunistic non-IFRS reporting is more prevalent in countries with more efficient law enforcement and investor protection. In such countries also excluding recurring items out of the non-IFRS earnings was seen to be more common. The same study considered Finland among the more developed countries when it comes to these institutional and economic factors. This would suggest that the nature of the Finnish reporting environment would not be the reason for the better-than-expected quality of the non-IFRS earnings measures. One possible explanation is also the nature of the companies included in the sample. Typically, only the very largest listed companies are looked at in the related literature. Since the sample here was derived from all Nasdaq Helsinki listed companies, the sample featured smaller companies than what are usually examined. Smaller listed companies may be characterized with different market conditions such as different target

audience for the financial reporting, different availability of financing and so on. Such characteristics and their effects on non-IFRS reporting could serve as another potential future research topic.

Hypothesis 2 expected that the ESMA Guidelines would have increased the quality of non-IFRS adjustments. Prior results from the US had suggested that such effect can be achieved through regulation (Jennings & Marques 2011; Black et al. 2012). The ESMA Guidelines include requirements such as that the naming of an APM must be consistent with its use and that the APMs should be reconciled to the closest IFRS line items. Ex ante such requirements would be expected to increase the quality of APMs, especially when measuring the quality through persistence of the non-IFRS adjustments.

The results however did not support there being any positive effect from the implementation of the ESMA Guidelines. The estimated coefficient for the effect of the ESMA Guidelines on reporting quality did not statistically differ from zero. The related coefficients were negative and had higher absolute values than the coefficients for quality of the adjustment items. Thus, while the ESMA Guidelines are not seen to have a statistically significant effect, the results are more aligned with the regulation having a negative effect on the quality of non-IFRS adjustments, not a positive one. The quantile regression model indicated that at higher percentiles of future performance, the negative effect of the ESMA Guidelines is actually pronounced and rather significant.

The most important explanation for the ESMA Guidelines having no effect on the quality of the non-IFRS adjustments is that, on average, the non-IFRS adjustments were already seen to be of high quality. When the typical non-IFRS adjustment is not found to be persistent and of low quality, then no room remains for the Guidelines to have any significant effect on the quality of these adjustments. No prior results focusing on Finland exist, but the results of this thesis suggest that over most of the past decade the quality of listed Finnish companies' non-IFRS earnings disclosures on average has been good. In this case it could be aligned with expectations that the regulation did not further improve the quality.

The reason for the Guidelines' lack of impact on reporting quality could also be related to the regulation itself. The ESMA Guidelines however are mostly similar in nature to the regulation introduced in the US in early 2000s, with both including requirements on issues such as reconciliation, naming, and presentation of APMs. If the content of the regulation is assumed to be effective, lack of enforcement could still water down the positive effect. In a report published at the end of 2019, ESMA found that only a small

minority of reporting companies complied with all principles of the Guidelines, though most companies did not completely ignore the requirements either (ESMA 2019). The results of the report imply that more work is needed before the possible benefits of the ESMA Guidelines are fully realized, a conclusion that could be supported by the results of this thesis.

It is also possible that the Guidelines have had effects on other variables not in the scope of this study. For one, the regulation is for the whole Europe, and therefore any decisive evaluation of the effects of the Guidelines should look at a more comprehensive sample of companies from different European countries. Secondly, even in a case where the non-IFRS adjustments were already of high quality, the ESMA Guidelines may have had other positive impacts, such as how much investors are able to trust and utilize these measures and whether the overall presentation has become clearer and more understandable for the end-users. As presented in chapter 2, analysts are very much open to expanded use of APMs as long as the use of these figures is also better regulated (CFA Institute 2018). Future research could further evaluate what kind of effects the Guidelines may have had for how analysts or investors actually use these measures.

Finally, hypothesis 3 predicted that high levels of impression management would be linked to non-IFRS adjustments of lower quality. Typically, firms could be presenting low quality non-IFRS figures when they are trying to mislead investors and overstate the true performance during the period. In such case the management is incentivized to follow such approach holistically in their financial communications: not just in presenting a single overstated figure but adjusting the surrounding message as well. Overly positive communications strategy can also be considered to act as something that conceals the misleading use of an APM. Prior results from the study by Guillamon-Saorin et al. (2017) supported such expectations.

Based on the results a clear association between high impression management and lower quality non-IFRS adjustments was identified. The related regression coefficient was in absolute terms higher than the coefficient for the other main variables in the model. The result is comparable to what Guillamon-Saorin et al. (2017) found. Interesting finding then is that even when the non-IFRS adjustments on average were found to be of good quality, when linked with high impression management the quality dropped significantly.

Several conclusions could be drawn from this. First, this is very compelling evidence for the validity of linking impression management and misleading reporting together. The hypothesis holds even in the reporting environment that overall is seen to be of high

quality. Companies are seen to engage in aggressive communications to inflict strong positive bias; even to mislead investors. Relatedly, results also further validate the employed impression management score, that is based on the framework from Brennan et al. (2009). Continuing to utilize the measure in future research could open new opportunities also outside the non-GAAP literature. Thirdly, the result can be linked to the existing research showing somewhat mixed results on non-GAAP or non-IFRS reporting. Since high impression management was also here linked to lower quality disclosures, it offers new proof that impression management is one important factor or determinant for making a distinction between useful and opportunistic non-GAAP reporting.

That also leads back to the objectives of the thesis. The first main objective was to examine the quality of non-IFRS figures and especially gain new evidence on how certain variables are linked to the quality of the adjustments. Because prior results are indecisive on the usefulness versus misuse of APMs, a logical step is to search for determinants that show when APMs can be useful and when less so. It was seen that among the Finnish sample, the quality of non-IFRS earnings reporting is generally good. On the other hand, the results of this thesis support the idea that high levels of impression management are linked to low quality non-IFRS reporting, even when outside of that no misuse is detected. The second main objective was to evaluate the effect of the ESMA Guidelines on quality of non-IFRS disclosures. Here the results showed that the regulation would not have had any real effect – but since the reporting quality was good already pre-2016 among the Finnish sample, any significant effect would not be expect.

Overall, this thesis provided new evidence on the Finnish non-IFRS reporting landscape, showing that during most of the 2010s the non-IFRS figures have been of good quality. On average the listed companies are not seen to be using the figures in a misleading way. When engaging in high impression management with the non-IFRS reporting, issues start to appear. Aggressive communications is found to be linked to lower quality non-IFRS disclosures. This suggest that in these cases companies are presenting misleading APMs and trying to inflict positive bias also in communications surrounding these measures. The findings have implications for the users of financial information especially when evaluating to what extent the discretionary performance measures can be relied on. For regulators it is worth considering what it means that through several linguistic tactics companies are trying to hide the misleading reporting. Especially so when the results suggest that the ESMA Guidelines, as currently in place, have not had any effect on the quality of the non-IFRS earnings reporting. To conclude, all capital market participants

can benefit from identifying and paying attention to overly aggressive impression management and financial communications, while especially for regulators the results suggest that holistic evaluation on effectiveness of the ESMA Guidelines on a European level may be warranted.

6.2 Limitations of the thesis and suggestions for future research

The goal of this thesis has been to examine non-IFRS earnings reporting and its quality in the sample of Finnish stock-listed companies, partly conditional to the ESMA Guidelines and impression management. The employed empirical approach and methods have certain limitations that are discussed next, after which suggestions for further research are proposed based both on the results and identified limitations of the thesis at hand.

The sample for the empirical tests was collected completely from Finnish stock-listed companies. The initial sample contained 976 firm-year observations for the time period of 2012 to 2018, while the final sample size was 343 firm-year observations. The initial sample is already limited by the sized of the Finnish stock market, which consists of roughly 120 constituents – before any exclusions. The sample size was further decreased due to excluding financial companies, observations with missing data, and observations not containing any non-IFRS figures. In non-GAAP research, the sample sizes are generally somewhat limited, since data on reported non-GAAP figures is not available from any central repository, rather it must be hand collected. Certain studies have used analyst non-GAAP earnings as proxies for reported non-GAAP figures to achieve bigger samples, but as discussed in chapter 4.1, this is not an especially good proxy. The manual coding of the impression management score also limits the possible sample size. Smaller sample size can be problematic for generalizing the results to a wider population of companies, but also for the validity of employed statistical models. While a sample of over 300 observations is plenty for statistical analysis and interference, it is possible that even larger sample would be needed to achieve needed power and degrees of freedom to properly employ the asymptotic properties of OLS.

Limiting the sample to Finnish companies limits the ability to generalize the results to a wider European or global context. This is especially relevant for evaluating the effects of the ESMA Guidelines. While the Guidelines are not here found to be effective here, the regulation may have significant effects among the population of all European listed companies. The sample can still be considered representative of the Finnish setting, with

over half of all of the listed companies being included in the scope in each year of the sample.

It is also important to consider possible effects of self-selection. Since the thesis was partly focused on effects of impression management related to the non-IFRS reporting, all observations with no non-IFRS figures were scoped out. Analyzing only firms reporting non-IFRS figures does not pay attention to the fact that certain variables or conditions may be affecting the decision to disclose these figures. The objectives of the thesis are directly linked to non-IFRS reporting, and the validity of the results are not considered to be affected by self-selection. Issue of self-selection is more relevant in that the results or conclusions may not be generalizable to companies not reporting non-IFRS information.

The sample contains three years of observations from the post-ESMA Guidelines period. This may not be enough for any decisive conclusions, even if the sample size is sufficient. More research is needed to identify whether the regulation reaches its target of improving the usability of APMs. Also, the employed measure of non-IFRS reporting quality is not the only way to evaluate effects of the regulation. Especially since the non-IFRS adjustments did not have significant persistence even prior to the Guidelines, examining the quality of non-IFRS disclosures through other measures such as information content could provide different results. The persistence of the non-IFRS adjustments could also be evaluated over longer time periods than a single year. Evaluating the persistence over a period of several years could lead to different results than what is seen here.

Several suggestions for future research could therefore be made. New studies on non-GAAP reporting should move on from simple evaluation of quality or usefulness of the adjusted figures. Rather the focus should be on identifying the variables or characteristics linked to usefulness or misuse of APMs, such as effects of board composition, compensation, and so on. Such results can help enhance the users of financial information to identify when the non-GAAP reporting might be valuable resource as opposed to misleading information. This also helps regulators improve the regulation and thus the functioning of the capital markets.

Since previous studies focus especially on the largest stock-listed companies, another important research area could be to evaluate the smaller listed companies' non-GAAP reporting. The market environment of smaller companies can be quite different to their large counterparts, due to dynamics such as investor characteristics, liquidity, and the

amount of resources available for financial reporting. It may be that studying only the largest companies' non-IFRS reporting gives only a limited picture of the whole truth.

Finally, more research on the effects of the ESMA Guidelines is needed. The effects should be considered on a wider sample of European companies in order to get more comprehensive results. Other measures of reporting quality could also be employed to examine whether the ESMA Guidelines have had other effects, for example on investors' confidence and trust towards the adjusted figures.

7 SUMMARY

While non-GAAP reporting has continuously increased in popularity among public companies over the past two decades, the research on the value of these performance measures has been somewhat mixed. These adjusted figures aim to improve the normal performance figures by diverging from what the accounting standards require. While a strong case can be made for considering earnings the ultimate performance metric of a firm, skepticism towards accounting earnings is not a new phenomenon. Based on the existing literature, both theoretical and practical support can be seen to exist for the idea that an adjusted non-GAAP earnings figure could be more valuable than the regulation-based counterpart. Research on non-GAAP reporting has also provided direct evidence that investors use non-GAAP figures, that the adjusted figures provide incremental information over the GAAP figures, and that the adjusted figures have predictive power over future performance.

Despite the potential value and benefits of non-GAAP reporting, lots of criticism has been directed toward these measures, which historically have been regulated only in a limited capacity. When companies are moving from regulated performance measures to figures based purely on their own definitions, a possibility of deception and fraudulent reporting exists. Previous studies have shown that non-GAAP figures almost always exceed the GAAP-counterpart, that the figures are used inconsistently and against the stated principles, also that the reporting is more common when firms are missing external benchmarks or when management is selling the shares after the earnings release. It can be seen that the existing results on usefulness and misuse of non-GAAP figures are somewhat mixed.

This thesis aimed to provide new evidence on the quality of non-GAAP reporting among Finnish listed companies, both overall and conditional to common European regulation and impression management. New contribution to the existing, mixed results can be gained by examining additional variables that are linked with the quality and misuse of these adjusted metrics. Linking together the overall financial communications around the non-GAAP figures allows for more detailed and comprehensive analysis of the reporting. Meanwhile new results are also provided by evaluating what kind of effects the ESMA Guidelines on alternative performance measures may have had on non-IFRS reporting after their implementation in 2016.

Overall, the results of the thesis indicated that the reported non-IFRS earnings figures were of good quality and that they were used in accordance with the stated definitions. The fact that the non-IFRS earnings adjustments did not have predictive power for either of future earnings or future cash flows increases the validity of the result. This finding differs from previous results, as the prior evidence generally suggests that the non-GAAP exclusions do predict future performance. Several factors may explain the difference to previous results, including differences in sample characteristics such as time period, evaluated country, and firm characteristics. Most of the previous results cover time period in the 2000s as opposed to 2010s, while the Finnish companies or smaller stock markets in general have not been the focus of any previous studies. Especially examining the role of firm characteristics in relation to non-GAAP reporting could offer several future research topics.

While the non-IFRS reporting overall was seen to be of good quality, this was not the case when the adjusted figure is accompanied with high levels of impression management. When companies are engaging in aggressive impression management around the non-IFRS figure the adjustments are seen to be of low quality, with the adjustments having significant predictive power over future performance. This result holds with both operationalizations of future performance. The result can be explained by firms employing more aggressive communication tactics when reporting APMs that are of lower quality: this way the message gains more emphasis, and the true performance may remain obfuscated. This result further validates the impression management score and results of Guillamon-Saorin et al. (2017): the previous results and the applicability of the score apply also in a new decade, in a new setting, and even when the reporting quality overall was good.

The ESMA Guidelines were not seen to have any effect on the quality of non-IFRS reporting. While in the US research has shown at least short-term benefits from the regulation, here the results are actually more aligned with the Guidelines having had a negative effect, though statistically no significant effect was in place. The most obvious explanation is that since the reporting quality overall was already good, there was no obvious room or need for improvements. Worth noting is also that examining only the Finnish setting is not enough to declare that the ESMA Guidelines have not had any effects at all. Similarly, the applied proxy of reporting quality may not capture fully the impact of the regulation, and more examination could be done on whether the regulation has increased

the stakeholders trust in the adjusted figures and whether they are actually used more often after the rules came into place.

The results overall show how the quality of non-IFRS measures can be seen to be good among the Finnish sample, even with the popularity of these figures still increasing. This means that on average the adjusted figures represent what they are stated to represent and could provide valuable information for all stakeholders of the companies. The users of financial information however should be careful when companies are engaging in aggressive communications in accompanying the non-IFRS figures with high levels of impression management. In such cases the quality of the adjustments is poor, and the exclusions made actually do not represent non-recurring items: applying skepticism is warranted. For regulators it might be worthwhile to consider the implications of the above results. Currently the ESMA Guidelines have not increased the quality of the non-IFRS disclosures but taking action against overly biased communications could provide beneficial results – either through better regulation or better enforcement of the existing rules.

REFERENCES

Abernathy, John L – Beyer, Brooke – Rapley, Eric T (2014) Earnings Management Constraints and Classification Shifting. *Journal of Business Finance & Accounting*, Vol. 41 (5), 600–626.

Bai, Xuelian – Dong, Yi – Hu, Nan (2019) Financial Report Readability and Stock Return Synchronicity. *Applied Economics*, Vol. 51 (4), 346–363.

Ball, Ray – Brown, Philip (1968) An Empirical Evaluation of Accounting Income Numbers. *Journal of Accounting Research*, Vol. 6 (2), 159–178.

Ball, Ray – Foster, George (1982) Corporate Financial Reporting: A Methodological Review of Empirical Research. *Journal of Accounting Research*, Vol. 20, 161–234.

Bansal, Naresh – Seetharaman, Ananth – Wang, Xu (2013) Managerial Risk-Taking Incentives and Non-GAAP Earnings Disclosures. *Journal of Contemporary Accounting & Economics*, Vol. 9 (1), 100–121.

Barth, Mary E – Cram, Donald P – Nelson, Karen K (2001) Accruals and the Prediction of Future Cash Flows. *The Accounting Review*, Vol. 76 (1), 27–58.

Bartov, Eli – Givoly, Dan – Hayn, Carla (2002) The Rewards to Meeting or Beating Earnings Expectations. *Journal of Accounting and Economics*, Vol. 33 (2), 173–204.

Beattie, Vivien (2014) Accounting Narratives and the Narrative Turn in Accounting Research: Issues, Theory, Methodology, Methods and a Research Framework. *The British Accounting Review*, Vol. 46 (2), 111–134.

Beaver, William H (1968) The Information Content of Annual Earnings Announcements. *Journal of Accounting Research*, Vol. 6, 67–92.

Beaver, William H – McNichols, Maureen F – Wang, Zach Z (2018) The Information Content of Earnings Announcements: New Insights from Intertemporal and Cross-Sectional Behavior. *Review of Accounting Studies*, Vol. 23 (1), 95–135.

Bentley, Jermiah W – Christensen, Theodore E – Gee, Kurt H – Whipple, Benjamin C (2018) Disentangling Managers' and Analysts' Non-GAAP Reporting. *Journal of Accounting Research*, Vol. 56 (4), 1039–1081.

Beyerlein, Andreas (2014) Quantile Regression - Opportunities and Challenges From a User's Perspective. *American Journal of Epidemiology*, Vol. 180 (3), 330–331.

Bhattacharya, Nilabhra – Black, Ervin L – Christensen, Theodore E – Larson, Chad R (2003) Assessing the Relative Informativeness and Permanence of Pro Forma

Earnings and GAAP Operating Earnings. *Journal of Accounting and Economics*, Vol. 36 (1), 285–319.

Bhattacharya, Nilabhra – Black, Ervin L – Christensen, Theodore E – Mergenthaler, Richard D (2007) Who Trades on Pro Forma Earnings Information? *The Accounting Review*, Vol. 82 (3), 581–619.

Black, Dirk E – Black, Ervin L – Christensen, Theodore E – Heninger, William G (2012) Has the Regulation of Pro Forma Reporting in the US Changed Investors' Perceptions of Pro Forma Earnings Disclosures? *Journal of Business Finance & Accounting*, Vol. 39 (7), 876–904.

Black, Dirk E – Christensen, Theodore E (2009) US Managers' Use of "Pro Forma" Adjustments to Meet Strategic Earnings Targets. *Journal of Business Finance & Accounting*, Vol. 36 (3), 297–326.

Black, Dirk E – Christensen, Theodore E (2018) Policy Implications of Research on Non-GAAP Reporting. *Research in Accounting Regulation*, Vol. 30 (1), 1–7.

Black, Dirk E – Christensen, Theodore E – Ciesielski, Jack T – Whipple, Benjamin C (2018) Non-GAAP Reporting: Evidence from Academia and Current Practice. *Journal of Business Finance & Accounting*, Vol. 45 (3), 259–294.

Black, Ervin L – Christensen, Theodore E – Kiosse, Paraskevi Vicky – Steffen, Thomas D (2015) Has the Regulation of Non-GAAP Disclosures Influenced Managers' Use of Aggressive Earnings Exclusions? *Journal of Accounting, Auditing & Finance*, Vol. 32 (2), 209–240.

Bond, David – Czernkowski, Robert – Lee, Yong-Suk – Loyeung, Anna (2017) Market Reaction to Non-GAAP Earnings around SEC Regulation. *Journal of Contemporary Accounting & Economics*, Vol. 13 (3), 193–208.

Bradshaw, Mark T – Sloan, Richard G (2002) GAAP versus The Street: An Empirical Assessment of Two Alternative Definitions of Earnings. *Journal of Accounting Research*, Vol. 40 (1), 41–66.

Brennan, Niamh M. – Guillamon-Saorin, Encarna – Pierce, Aileen (2009) Methodological Insights: Impression Management: Developing and Illustrating a Scheme of Analysis for Narrative Disclosures – a Methodological Note. *Accounting, Auditing & Accountability Journal*, Vol. 22 (5), 789–832.

Brockman, Paul – Martin, Xiumin – Puckett, Andy (2010) Voluntary Disclosures and the Exercise of CEO Stock Options. *Journal of Corporate Finance*, Vol. 16 (1), 120–136.

Brown, Lawrence D – Call, Andrew C – Clement, Michael B – Sharp, Nathan Y (2015) Inside the "Black Box" of Sell-Side Financial Analysts. *Journal of Accounting Research*, Vol. 53 (1), 1–47.

Brown, Lawrence D – Caylor, Marcus L (2005) A Temporal Analysis of Quarterly Earnings Thresholds: Propensities and Valuation Consequences. *The Accounting Review*, Vol. 80 (2), 423–440.

Brown, Lawrence D – Higgins, Huong N (2005) Managers' Forecast Guidance of Analysts: International Evidence. *Journal of Accounting and Public Policy*, Vol. 24 (4), 280–299.

Caglio, Ariela – Melloni, Gaia – Perego, Paolo (2019) Informational Content and Assurance of Textual Disclosures: Evidence on Integrated Reporting. *The European Accounting Review*, Vol. 29 (1), 1–29.

Cameron, Robyn – Percy, Majella – Stevenson-Clarke, Peta (2012) Do Large Australian Companies Emphasise Non-GAAP Financial Measures: Over Statutory Net Profit (GAAP) in Annual Reports? *Journal of the Securities Institute of Australia*, Vol. 3 (1), 19–25.

CFA Institute (2016) *Investor Uses, Expectations, and Concerns on Non-GAAP Financial Measures*. <https://www.cfainstitute.org/-/media/documents/support/advocacy/investor-uses-expectations-concerns-on-non-gaap.ashx>, accessed 5.12.2020.

CFA Institute (2018) *Member Survey Report: Alternative Performance Measures — The Latest on Investor Expectations: The Way Forward on APMs*. <https://www.cfainstitute.org/en/research/survey-reports/alternative-performance-measures>, accessed 10.12.2020.

CFA Society UK (2015) *Non-IFRS Earnings and Alternative Performance Measures: Ensuring a Level Playing Field*. <https://www.cfauk.org/professionalism/research-and-position-papers>, accessed 1.1.2019.

Chambers, R. J. (1993) Positive Accounting Theory and the PA Cult. *Abacus*, Vol. 29 (1), 1–26.

Christensen, Theodore E – Drake, Michael S – Thornock, Jacob R (2014) Optimistic Reporting and Pessimistic Investing: Do Pro Forma Earnings Disclosures Attract Short Sellers? *Contemporary Accounting Research*, Vol. 31 (1), 67–102.

Christenson, Charles (1983) The Methodology of Positive Accounting. *The Accounting Review*, Vol. 58 (1), 1–22.

Ciesielski, Jack T – Henry, Elaine (2017) Accounting's Tower of Babel: Key Considerations in Assessing Non-GAAP Earnings. *Financial Analysts Journal*, Vol. 73 (2), 34–50.

Cooper, Kerry – Keim, Gerald D (1983) The Economic Rationale for the Nature and Extent of Corporate Financial Disclosure Regulation: A Critical Assessment. *Journal of Accounting and Public Policy*, Vol. 2 (3), 189–205.

Curtis, Asher B – McVay, Sarah E – Whipple, Benjamin C (2013) The Disclosure of Non-GAAP Earnings Information in the Presence of Transitory Gains. *The Accounting Review*, Vol. 89 (3), 933–958.

Dechow, Patricia – Ge, Weili – Schrand, Catherine (2010) Understanding Earnings Quality: A Review of the Proxies, Their Determinants and Their Consequences. *Journal of Accounting & Economics*, Vol. 50 (2–3), 344–401.

Dechow, Patricia M (1994) Accounting Earnings and Cash Flows as Measures of Firm Performance: The Role of Accounting Accruals. *Journal of Accounting and Economics*, Vol. 18 (1), 3–42.

Dichev, Ilia D – Tang, Wei Vicki (2008) Matching and the Changing Properties of Accounting Earnings over the Last 40 Years. *The Accounting Review*, Vol. 83 (6), 1425–1460.

Distante, Roberta – Petrella, Ivan – Santoro, Emiliano (2018) Gibrat's Law and Quantile Regressions: An Application to Firm Growth. *Economics letters*, Vol. 164, 5–9.

Doyle, Jeffrey T – Jennings, Jared N – Soliman, Mark T (2013) Do Managers Define Non-GAAP Earnings to Meet or Beat Analyst Forecasts? *Journal of Accounting and Economics*, Vol. 56 (1), 40–56.

Doyle, Jeffrey T – Lundholm, Russell J – Soliman, Mark T (2003) The Predictive Value of Expenses Excluded from Pro Forma Earnings. *Review of Accounting Studies*, Vol. 8 (2), 145–174.

Elliott, W Brooke (2006) Are Investors Influenced by Pro Forma Emphasis and Reconciliations in Earnings Announcements? *The Accounting Review*, Vol. 81 (1), 113–133.

Entwistle, Gary – Feltham, Glenn – Mbagwu, Chima (2012) Credibility Attributes and Investor Perceptions of Non-GAAP Earnings Exclusions. *Accounting Perspectives*, Vol. 11 (4), 229–257.

Entwistle, Gary – Feltham, Glenn – Mbagwu, Chima (2006) Financial Reporting Regulation and the Reporting of pro Forma Earnings. *Accounting Horizons*, Vol. 20 (1), 39–55.

ESMA (2015) *ESMA Publishes Guidelines for Issuers Performance Measures*. Press release. Paris.

ESMA (2017) *Questions and Answers*. <https://www.esma.europa.eu/press-news/esma-news/esma-publishes-new-qa-alternative-performance-measures-guidelines>, accessed 1.11.2019.

ESMA (2019) *Report On the Use of Alternative Performance Measures and on the Compliance with ESMA's APM Guidelines*. <https://www.esma.europa.eu/press-news/esma-news/esma-eu-issuers-need-improve-their-disclosure-alternative-performance-measures>, accessed 5.12.2020.

ESMA Guidelines on Alternative Performance Measures (2015/1057). The European Securities and Markets Authority (ESMA).

Fan, Yun – Barua, Abhijit – Cready, William M – Thomas, Wayne B (2010) Managing Earnings Using Classification Shifting: Evidence from Quarterly Special Items. *The Accounting Review*, Vol. 85 (4), 1303–1323.

Fiechter, Peter (2013) The Effects of Compensation and Board Quality on Non-GAAP Disclosures in Europe. *The International Journal of Accounting*, Vol. 48 (3), 318–323.

FIN-FSA (2017) *Market Newsletter 4/2017*. <https://helda.helsinki.fi/bof/bitstream/handle/123456789/15333/Market_newsletter_4_2017.pdf?sequence=1&isAllowed=y>, accessed 8.9.2019.

FIN-FSA (2019) *Market Newsletter 1/2019*. <https://helda.helsinki.fi/bof/bitstream/handle/123456789/16100/market_newsletter_1_2019.pdf?sequence=1&isAllowed=y>, accessed 8.9.2019.

Frankel, Richard – McVay, Sarah – Soliman, Mark (2011) Non-GAAP Earnings and Board Independence. *Review of Accounting Studies*, Vol. 16 (4), 719–744.

García Osma, Beatriz – Guillamón-Saorín, Encarna (2011) Corporate Governance and Impression Management in Annual Results Press Releases. *Accounting, Organizations and Society*, Vol. 36 (4–5), 187–208.

Graham, John R – Harvey, Campbell R – Rajgopal, Shiva (2005) The Economic Implications of Corporate Financial Reporting. *Journal of Accounting and Economics*, Vol. 40 (1), 3–73.

Guillamon-Saorin, Encarna – Isidro, Helena – Marques, Ana (2017) Impression Management and Non-GAAP Disclosure in Earnings Announcements. *Journal of Business Finance & Accounting*, Vol. 44 (3), 448–479.

Hayes, Andrew F – Cai, Li (2007) Using Heteroskedasticity-Consistent Standard Error Estimators in OLS Regression: An Introduction and Software Implementation. *Behavior Research Methods*, Vol. 39 (4), 709–722.

Hinkley, David V (1977) Jackknifing in Unbalanced Situations. *Technometrics*, Vol. 19 (3), 285–292.

Holthausen, Robert W (1990) Accounting Method Choice: Opportunistic Behavior, Efficient Contracting, and Information Perspectives. *Journal of Accounting and Economics*, Vol. 12 (1), 207–218.

Hoogervorst, Hans (31.3.2015) *Mind the Gap (Between Non-GAAP and GAAP)*. <http://thegaap.net/mind-the-gap-between-non-gaap-and-gaapspeech-by-hans-hoogervorst/>, accessed 1.11.2019.

Hoogervorst, Hans (11.5.2016) *Performance Reporting and the Pitfalls of Non-GAAP Metrics*. <https://www.ifrs.org/-/media/feature/news/speeches/2016/hans-hoogervorst-eaa-annual-conference-may-2016.pdf>, accessed 1.11.2019.

IFRS Foundation (2019) *The Annotated IFRS® Standards: Standards Required 1 January 2019*. 1st ed. London: IFRS Foundation.

Isidro, Helena – Marques, Ana (2013) The Effects of Compensation and Board Quality on Non-GAAP Disclosures in Europe. *The International Journal of Accounting*, Vol. 48 (3), 289–317.

Isidro, Helena – Marques, Ana (2015) The Role of Institutional and Economic Factors in the Strategic Use of Non-GAAP Disclosures to Beat Earnings Benchmarks. *European Accounting Review*, Vol. 24 (1), 95–128.

Jennings, Ross – Marques, Ana (2011) The Joint Effects of Corporate Governance and Regulation on the Disclosure of Manager-Adjusted Non-GAAP Earnings in the US. *Journal of Business Finance & Accounting*, Vol. 38 (3), 364–394.

Jensen, Michael C (1983) Organization Theory and Methodology. *The Accounting Review*, Vol. 58 (2), 319–339.

Jensen, Michael C – Meckling, William H (1976) Theory of the Firm: Managerial Behavior, Agency Costs and Ownership Structure. *Journal of Financial Economics*, Vol. 3 (4), 305–360.

Jordan, Charles E – Waldron, Marilyn A – Clark, Stanley J (2011) An Analysis Of The Comparative Predictive Abilities Of Operating Cash Flows, Earnings, And Sales. *Journal of Applied Business Research*, Vol. 23 (3).

Jorgensen, Bjorn – Li, Jing – Sadka, Gil (2012) Earnings Dispersion and Aggregate Stock Returns. *Journal of Accounting and Economics*, Vol. 53 (1), 1–20.

Kaplan, Steven E – Ruland, Robert G (1991) Positive Theory, Rationality and Accounting Regulation. *Critical Perspectives On Accounting*, Vol. 2 (4), 361–374.

Kasznik, Ron – McNichols, Maureen F (2002) Does Meeting Earnings Expectations Matter? Evidence from Analyst Forecast Revisions and Share Prices. *Journal of Accounting Research*, Vol. 40 (3), 727–759.

Kim, Myungsun – Kross, William (2005) The Ability of Earnings to Predict Future Operating Cash Flows Has Been Increasing-Not Decreasing. *Journal of Accounting Research*, Vol. 43 (5), 753–780.

Koenker, Roger – Bassett, Gilbert (1978) Regression Quantiles. *Econometrica*, Vol. 46 (1), 33–50.

Koenker, Roger – Hallock, Kevin F (2001) Quantile Regression. *The Journal of Economic Perspectives*, Vol. 15 (4), 143–156.

Kothari, S P (2001) Capital Markets Research in Accounting. *Journal of Accounting & Economics*, Vol. 31 (1–3), 105–231.

Kothari, S P – Ramanna, Karthik – Skinner, Douglas J (2010) Implications for GAAP from an Analysis of Positive Research in Accounting. *Journal of Accounting & Economics*, Vol. 50 (2–3), 246–286.

Leary, Mark R – Kowalski, Robin M (1990) Impression Management: A Literature Review and Two-Component Model. *Psychological Bulletin*, Vol. 107 (1), 34–47.

Lehavy, Reuven – Li, Feng – Merkley, Kenneth (2011) The Effect of Annual Report Readability on Analyst Following and the Properties of Their Earnings Forecasts. *The Accounting Review*, Vol. 86 (3), 1087–1115.

Leung, Edith – Veenman, David (2018) Non-GAAP Earnings Disclosure in Loss Firms. *Journal of Accounting Research*, Vol. 56 (4), 1083–1137.

Llewelyn, Sue (2003) What Counts as "Theory" in Qualitative Management and Accounting Research? Introducing Five Levels of Theorizing. *Accounting, Auditing, & Accountability*, Vol. 16 (4), 662–708.

Long, J Scott – Ervin, Laurie H (2000) Using Heteroscedasticity Consistent Standard Errors in the Linear Regression Model. *The American Statistician*, Vol. 54 (3), 217–224.

Lougee, Barbara A – Marquardt, Carol A (2004) Earnings Informativeness and Strategic Disclosure: An Empirical Examination of "Pro Forma" Earnings. *The Accounting Review*, Vol. 79 (3), 769–795.

Loughran, Tim – McDonald, Bill (2011) When Is a Liability Not a Liability? Textual Analysis, Dictionaries, and 10-Ks. *The Journal of Finance (New York)*, Vol. 66 (1), 35–65.

Loughran, Tim – McDonald, Bill (2014) Measuring Readability in Financial Disclosures. *The Journal of Finance*, Vol. 69 (4), 1643–1671.

Marques, Ana (2006) SEC Interventions and the Frequency and Usefulness of Non-GAAP Financial Measures. *Review of Accounting Studies*, Vol. 11 (4), 549–574.

Marques, Ana (2017) Non-GAAP Earnings: International Overview and Suggestions for Future Research. *Meditari Accountancy Research*, Vol. 25 (3), 318–335.

Matsumoto, Dawn A (2002) Management's Incentives to Avoid Negative Earnings Surprises. *The Accounting Review*, Vol. 77 (3), 483–514.

Mian, Shehzad L – Smith, Clifford W (1990) Incentives for Unconsolidated Financial Reporting. *Journal of Accounting and Economics*, Vol. 12 (1), 141–171.

Michael, Howard (2019) Misuse of Non-Mandatory Earnings Reporting by Companies. *Meditari Accountancy Research*, Vol. 27 (1), 125–146.

Ribeiro, Andrea – Shan, Yaowen – Taylor, Stephen (2019) Non-GAAP Earnings and the Earnings Quality Trade-Off. *Abacus*, Vol. 55 (1), 6–41.

Richardson, Scott – Teoh, Siew Hong – Wysocki, Peter D (2004) The Walk-down to Beatable Analyst Forecasts: The Role of Equity Issuance and Insider Trading Incentives. *Contemporary Accounting Research*, Vol. 21 (4), 885–924.

Rogers, Jonathan L – Van Buskirk, Andrew – Zechman, Sarah L C (2011) Disclosure Tone and Shareholder Litigation. *The Accounting Review*, Vol. 86 (6), 2155–2183.

Shiah-Hou, Shin-Rong – Teng, Yi-Yun (2016) The Informativeness of Non-GAAP Earnings after Regulation G? *Finance Research Letters*, Vol. 18, 184–192.

Tinker, Anthony M – Merino, Barbara D – Neimark, Marilyn Dale (1982) The Normative Origins of Positive Theories: Ideology and Accounting Thought. *Accounting, Organizations and Society*, Vol. 7 (2), 167–200.

Venter, Elmar R – Emanuel, David – Cahan, Steven F (2014) The Value Relevance of Mandatory Non- GAAP Earnings. *Abacus*, Vol. 50 (1), 1–24.

Walker, Martin (2013) How Far Can We Trust Earnings Numbers? What Research Tells Us about Earnings Management. *Accounting and Business Research*, Vol. 43 (4), 445–481.

Watts, Ross L (1977) Corporate Financial Statements, A Product of the Market and Political Processes. *Australian Journal of Management*, Vol. 2 (1), 53–75.

Watts, Ross L – Zimmerman, Jerold L (1978) Towards a Positive Theory of the Determination of Accounting Standards. *The Accounting Review*, Vol. 53 (1), 112–134.

Watts, Ross L – Zimmerman, Jerold L (1979) The Demand for and Supply of Accounting Theories: The Market for Excuses. *The Accounting Review*, Vol. 54 (2), 273–305.

Watts, Ross L – Zimmerman, Jerold L (1990) Positive Accounting Theory: A Ten Year Perspective. *The Accounting Review*, Vol. 65 (1), 131–156.

Whipple, Benjamin C. (2015) The Great Unknown: Why Exclude Other Items from Non-GAAP Earnings Calculations in the Post-Reg G World? Working paper. <https://ssrn.com/abstract=2480663>, accessed 5.12.2020.

Whitley, R. D. (1988) The Possibility and Utility of Positive Accounting Theory. *Accounting, Organizations and Society*, Vol. 13 (6), 631–645.

Whittred, Greg (1987) The Derived Demand for Consolidated Financial Reporting. *Journal of Accounting and Economics*, Vol. 9 (3), 259–285.

Wooldridge, Jeffrey (2013) *Introductory Econometrics: A Modern Approach, Fifth Edition*. South-Western Cengage Learning, Mason, Ohio.