Disclosure Regulation, Governance Mechanisms and their Implementation in Practice

Empirical Evidence on
Critical Accounting Policies and Board Diversity

Dissertation

zur Erlangung des akademischen Grades eines
Doktors der Wirtschaftswissenschaften der
Rechts- und Wirtschaftswissenschaftlichen Fakultät
der Universität Bayreuth

vorgelegt von
Hendrik Rupertus
aus Weinheim

Dekan:	Prof. Dr. Jörg Gundel
Erstberichterstatter:	Prof. Dr. Rolf Uwe Fülbier, StB
Zweitberichterstatter:	Prof. Dr. Friedrich Sommer
Tag der mündlichen Prüfung:	17.06.2019

Danksagung

Die vorliegende Arbeit entstand während meiner Tätigkeit als wissenschaftlicher Mitarbeiter am Lehrstuhl für Internationale Rechnungslegung von Herrn Prof. Dr. Rolf Uwe Fülbier an der Universität Bayreuth. Sie wurde im Sommersemester 2019 von der Rechts- und Wirtschaftswissenschaftlichen Fakultät der Universität Bayreuth als Dissertation angenommen.

Der erfolgreiche Abschluss dieses „Großprojekts“ wäre ohne die Unterstützung einer Vielzahl an Personen nicht möglich gewesen. Mein größter Dank gilt meinem Doktorvater Herrn Prof. Dr. Rolf Uwe Fülbier, der mir nicht nur die Möglichkeit einer Promotion gegeben, sondern auch meine Begeisterung und Freude an der (empirischen) Rechnungslegungsforschung entfacht und neue Blickwinkel auf die Internationale Rechnungslegung aufgezeigt hat. Ohne seine fortwährende Unterstützung, stetige Diskussionsbereitschaft und immerzu hilfreichen Anmerkungen würde es diese Arbeit nicht geben. Doch auch abseits des beruflichen Alltags haben mich viele seiner persönlichen und menschlichen Eigenschaften geprägt. Danke für diese unvergessliche und lehrreiche Zeit!

Herrn Prof. Dr. Friedrich Sommer danke ich für die freundliche Übernahme und Mühen des Zweitgutachtens sowie die spannende, konstruktive und unterhaltsame Zusammenarbeit, nicht nur im Rahmen meiner Dissertation. Herrn Prof. Dr. Klaus Schäfer gilt mein Dank für seinen Vorsitz der Prüfungskommission sowie für die Zusammenarbeit in den vergangenen Jahren.

Meinen (ehemaligen) Kolleginnen und Kollegen danke ich für eine unvergessliche gemeinsame Zeit am Lehrstuhl! Mit Marcus Bravidor, Thomas Loy, Christian Mehnert, Claudia Roberts, Christina Scharf, Christian Wittmann, Niko J. Wolf sowie den „Jungbullen“ Klara Lösse, Lorenz Piering und Jan Seitz verbinde ich wundervolle Erinnerungen, spannende Diskussionen sowie zahlreiche unterhaltsame Momente, auch abseits des Lehrstuhls. Dies gilt auch für Johanna Held, Benjamin Keller, Malte Klein, Maximilian Mayer, Alfred Nickisch, Claudia Schulze und Michael Popp. Vielen Dank auch an alle studentischen Hilfskräfte für die intensive Zusammenarbeit sowie die Unterstützung beim Gelingen meiner Promotion. Besonders hervorheben möchte ich hierbei Florian Federsel, Isabell Keller, Elisabeth Kuhn, Simon Lemnitzer, Marie-Thérèse Meyer, Anna Mollat und Johanna Paraknewitz.

Besonderer Dank gebührt meiner wunderbaren Familie, insbesondere meiner Mutter Astrid, meinem Vater Hans-Jürgen, meinen Schwestern Annika und Mareike, meinen Großeltern Elfriede, Erika und Jan und meinen baldigen Schwiegereltern Elfriede und Walter für ihre fortwährende Unterstützung und aufmunternden Worte. Insbesondere möchte ich Mareike für das Korrekturlesen meiner gesamten Arbeit danken! Zu guter Letzt danke ich meiner Freundin Laura, die mich nicht nur während der Promotionszeit „ertragen“, sondern auch an jedem einzelnen Tag bei meinem Vorhaben unterstützt und mir in anstrengenden Zeiten den Rücken freigehalten hat. Vielen Dank, dass es dich gibt!

Bayreuth, 02. Juli 2019 Hendrik Rupertus

Abstract

This dissertation presents four chapters on disclosure regulation and corporate governance mechanisms. While Chapter 1 and 2 focus on disclosures about highly uncertain accounting policies ('*critical accounting policies*', CAPs), Chapter 3 and 4 examine gender diversity on corporate boards.

The first chapter analyzes the regulatory framework, existing research, as well as implementation in practice of CAP disclosures. Using a sample of Standard & Poor's (S&P) 500 firms between 2001 and 2016 with hand-collected data about 35,686 CAPs, I provide initial evidence on the occurrence of CAPs over time and how related disclosures comply with SEC guidelines. In addition, I explore the reporting characteristics of CAPs by analyzing the length, textual similarity, specificity, and readability of respective disclosures. My results enhance current knowledge about the number, regulatory framework, and content of CAPs. In particular, it further allows indications about factors that determine a firm's decision to flag an accounting policy as critical, potential effects of CAPs on capital markets, and their usefulness for financial statement users.

The second chapter analyzes whether CAPs are useful to identify single financial statement positions that are highly uncertain. In accordance with prior literature, we assume that financial statement positions flagged as CAPs contain a higher degree of measurement uncertainties, and are thus less persistent with respect to future cash flows. Our results support this assumption. We further find that accrual components flagged as CAPs are in fact not less useful in predicting future cash flows *per se*. It additionally depends on the importance and to a certain extent on the specificity for a given firm. To best of our knowledge, there is no evidence on how to identify subjective and uncertain accruals that are less persistent with respect to future cash flows. We show that CAPs provide such information. Thus, we empirically demonstrate that CAPs represents a suitable channel to communicate measurement uncertainties.

The third chapter explores institutional supply- and demand-side factors associated with global differences in female board representation. The results show that functioning outside investor protection and a societal climate of gender equality contributes to more

women on boards, mainly by fostering the supply of suitable candidates. Our results reveal that longer-term supply side factors seem to be necessary to complement short-term demand-side regulation to increase female board representation.

The fourth chapter examines long-term effects of board gender diversity on capital markets. We analyze how stakeholders perceive female board members in the long-term. Further, we explore whether firms get punished by investors if they do not 'comply' with investors' and other stakeholders' expectations about gender diverse boards. Based on 8,872 firm-year observation from 13 countries, we conclude that investors seem to perceive female and male board members as being equivalent in the long-term and do not base their investment decisions on directors' gender. Moreover, we find no evidence for significant reduced or increased stock returns for firms that deviate from the expected ratio of female board members. While academic research claims that female board appointments may have short-term effects on capital markets, it seems that the market corrects this mispricing over the long run.

Table of Contents

List of Tables

List of Figures

INTRODUCTORY SUMMARY

Individual parties, such as share- and stakeholders, negotiate a set of agreements, obligations and rights with the corporation, which can be seen as a nexus of contractual relationships to reduce transaction and contracting costs (e.g., Alchian and Demsetz 1972; Coase 1937; Fama 1980; Fama and Jensen 1983; Jensen and Meckling 1976). Principal-agent theory elaborates on the relationship between two contractual parties, in which the principal (e.g., shareholder) delegates decision-making authority, responsibilities and work to the agent (e.g., management) (Jensen and Meckling 1976). Since most contracts are incomplete and not easily enforceable, problems between the agent and principal arise (Fama and Jensen 1983). Such conflicts lead to agency costs, including costs of writing, monitoring, structuring, and bonding a set of contracts (e.g., Fama and Jensen 1983; Jensen and Meckling 1976; Maassen 1999; Mallin 2016). Furthermore, in contrast to the principal, the agent has access to proprietary and superior information about a firm's current and future performance, resulting in information asymmetries.

In order to resolve the above-mentioned disadvantages, individuals contracting with a firm desire information and specific corporate structures that reduce agency costs, information asymmetries, and ensure satisfaction and compliance with the contractual terms and obligations (Bushman and Smith 2003; Healy and Palepu 2001). Agency theory views corporate disclosures as well as governance mechanisms as two potential channels through which principal-agent conflicts can be mitigated (Bushman and Smith 2003; Healy and Palepu 2001; Mallin 2016; Williamson 1984). On the one hand, the disclosure of relevant information enables principals to monitor contractual rights and duties, and to evaluate how the agent has allocated a firm's resources. On the other hand, governance mechanisms monitor and discipline the agents and ensure that they act in the interest of the principals (e.g., Healy and Palepu 2001).[1] Overall, in order to assess that a firm's management is compliant with the contractual arrangements, a firm has to be financially stable, well managed, and profitable in the future (e.g., Bushman and Smith 2003; Mallin 2016).

[1] For instance, investors and other shareholders need information and assurance about the efficiency and uncertainties of their financial investments. Other stakeholder groups, such as employees, suppliers, or customers are interested in a firm's ability to pay salaries, to secure future pension obligations, or to supply goods and services.

In practice, corporate disclosures, and especially financial reporting, can be defined as the communication of financial and non-financial information between insiders of a company (e.g., the management) and outsiders (e.g., investors[2]) (Healy and Palepu 2001). The firm provides its disclosures in the form of primary financial statements, notes to the financial statements, the Management, Discussion and Analysis (MD&A), and other regulatory filings (Barckow 2018; Healy and Palepu 2001).[3] Thereby, financial statement figures are supplemented by narrative information to enhance the understanding of investors about a firm's economic, business, and accounting environment (Palepu et al. 2016). According to the conceptual frameworks provided by the Financial Accounting Standards Board (FASB) as well as the International Accounting Standards Board (IASB), the main objective of financial reporting is to give a 'true and fair view' about the financial condition of the firm and to provide decision useful information to current and potential investors (FASB 2010; IASB 2018). Having said that, corporate disclosures shall have the ability to enhance a firm's information environment, reduce agency costs, and costs of external financing (Bushman and Smith 2003).[4] In this context, prior studies find that how and which financial accounting information are presented affects investors' assessment of future cash flows, mitigate information asymmetries, as well as the average cost of capital (e.g., Barth and Schipper 2008; Bushman and Smith 2003; Easley and O'Hara 2004; Lambert et al. 2012). Thus, financial transparency, defined as the overall availability of reliable and relevant firm-specific information, plays a key role in the context of corporate disclosures (Bushman et al. 2004).

Although financial reporting might be helpful to reduce information asymmetries between insiders and outsiders, corporate governance represents another channel to reduce agency conflicts (e.g., Mallin 2016; Shleifer and Vishny 1997; Tricker 2015; Williamson 1984). From an international point of view, the development of corporate governance has

[2] Corporate disclosures can also be directed to share- and stakeholders other than investors. However, investors are the primary users in the context of financial reporting (FASB 2010; IASB 2018). Consequently, I focus on investors thereafter.

[3] There are also other documents attached to financial reports such as the corporate governance reports, quarterly reports/statements, half year reports, and ad hoc announcements. There are also other channels to communicate information, such as management forecasts, conference calls, press releases, analyst meetings, and the internet (Healy and Palepu 2001).

[4] Bushman and Smith (2003) also argue that financial accounting information helps managers and investors in identifying 'good' and 'bad' projects to evaluate investment opportunities.

been affected by various disciplines (e.g., accounting, finance, economics, and organizational behaviour), economic theories[5], cultural and legally aspects, and other structural differences (Mallin 2016). Consequently, several definitions of corporate governance exist.[6] For example, Sir Adrian Cadbury's *Report on the Financial Aspects of Corporate Governanc*e defines corporate governance as a system in which companies are directed, managed, and controlled, while the board of directors is responsible for the governance of a company (Cadbury 1992). Shleifer and Vishny (1997) see corporate governance as a way how "suppliers of finance assure themselves of getting a return on investment" (Shleifer and Vishny 1997, p. 737). Blair (1995) describes corporate governance in the context of *who* controls public companies, *how* the control is carried out, and *what* companies are capable of. In particular, the board of directors represents a key role for reducing agency problems because it controls and leads a company. Moreover, it ensures that the management acts in the interest of share- and stakeholders (e.g., Healy and Palepu 2001) and represents the link between both parties (Mallin 2016; Monks and Minow 2011; Tricker 2015).

The board has several responsibilities that are essential for a good corporate governance. It formulates strategies, defines a company's direction, monitors the management, and executives activities, as well as provides accountability towards the shareholders (e.g., Burke 1997; Fondas 2000; Palepu et al. 2016). Moreover, the board can appoint subcommittee, such as the audit committee, remuneration committee, nomination committee, and ethics committee, enabling directors to meet independently from the board and delegating board activities to better focus on specific tasks (e.g., Mallin 2016; Tricker 2015). The composition of boards, in particular the diversity of the board members, has received great attention in the academic literature, international press (e.g., Grosvold et al. 2007; McGregor 2014; Olson 2019), and on the side of institutional investors, regulators, and other stakeholders (e.g., Byoun et al. 2016; Cao and Donnelly 2010; Coffey and Fryxell

[5] There exist several theories in explaining corporate governance, such as agency theory (Jensen and Meckling 1976; Fama 1980), transaction cost economics (Coase 1937; Williamson 1984), stakeholder theory (Jensen 2001; Mallin 2016), and stewardship theory (Donaldson and Davis 1991). However, I mainly refer to agency theory because it can be seen as one of the main theories associated with the development of corporate governance (e.g., Mallin 2016).

[6] According to Tricker (2015), corporate governance can be defined from five distinct perspectives: The operational perspective (Cadbury 1992), relationship perspective (Monks and Minow 2011), stakeholder perspective (Demb and Neubauer 1992), financial economics perspective (Shleifer and Vishny 1997), and social perspective (Blair 1995).

1991; SEC 2009). Board diversity should ensure a broad spectrum of members with distinct skills, experience, and knowledge relevant to a firm's business and industry. Further, diversity in terms of gender, age, race, and nationality should provide different perspectives on tasks of the board and develop new approaches to solutions and strategies. From a theoretical perspective, a balance of representatives with distinct contractual interests might strengthen the capability of boards and reflect the views of various stakeholder groups (Adams and Kirchmaier 2015; Mallin 2016; Tricker 2015).[7] To sum up, the board of directors and its composition is essential for an effective corporate governance (Fama and Jensen 1983; Mallin 2016).

Nonetheless, dubious governance practices and missing corporate disclosures have led to corporate scandals and failures at the beginning of the 21st century, such as Enron, WorldCom, Tyco, and the financial crisis (e.g., Healy and Palepu 2003; Mallin 2016). Consequently, financial transparency and corporate governance structures were questioned leading to a loss of confidence on capital markets. Since then, standard setters, regulators, practitioners, and the international press extensively discuss the improvement of corporate disclosures and governance mechanisms. In response to the scandals, various major regulatory changes have taken place in recent years. On the one hand, there has been an increasing concern that financial reporting requirements do not fulfil their intended purpose due to complex, unspecific, and opaque information. Thus, regulators revised and supplemented disclosure requirements to enhance the quality and transparency of financial disclosures to make capital markets more efficient. On the other hand, several countries and institutions introduced laws and revised their corporate governance codes[8] to sharpen the regulatory framework in the areas of leadership, management control, responsibilities, and board composition. Central to these debates is the board of directors, in

[7] There exist different types of board structures. The unitary board structure is predominant in the majority of European Union (EU) Member States and in the U.S. and consists of one single board with executive and non-executive directors. The unitary board is responsible for all tasks and activities of a company. Austria, Germany, Denmark, and the Netherlands have a dual board system, which is characterized by a supervisory board and an executive board of management. The shareholders appoint the members of the supervisory board and the supervisory board appoints the members of the management board. Whereas the supervisory board monitor the business and the activities of the management board, the management board runs the business (e.g., Mallin 2016; Monks and Minow 2011).

[8] Corporate governance codes have been issued by a variety of countries and institutions, such as the Sarbanes Oxley Act in the U.S. (U.S. Government 2002), the German Corporate Governance Code (Regierungskommission Deutscher Corporate Governance Kodex 2017), the U.K. Corporate Governance Code (Financial Reporting Council 2018), and Principles of Corporate Governance published by the Organisation for Economic Co-operation and Development (OECD) (Organisation for Economic Co-operation and Development 2015).

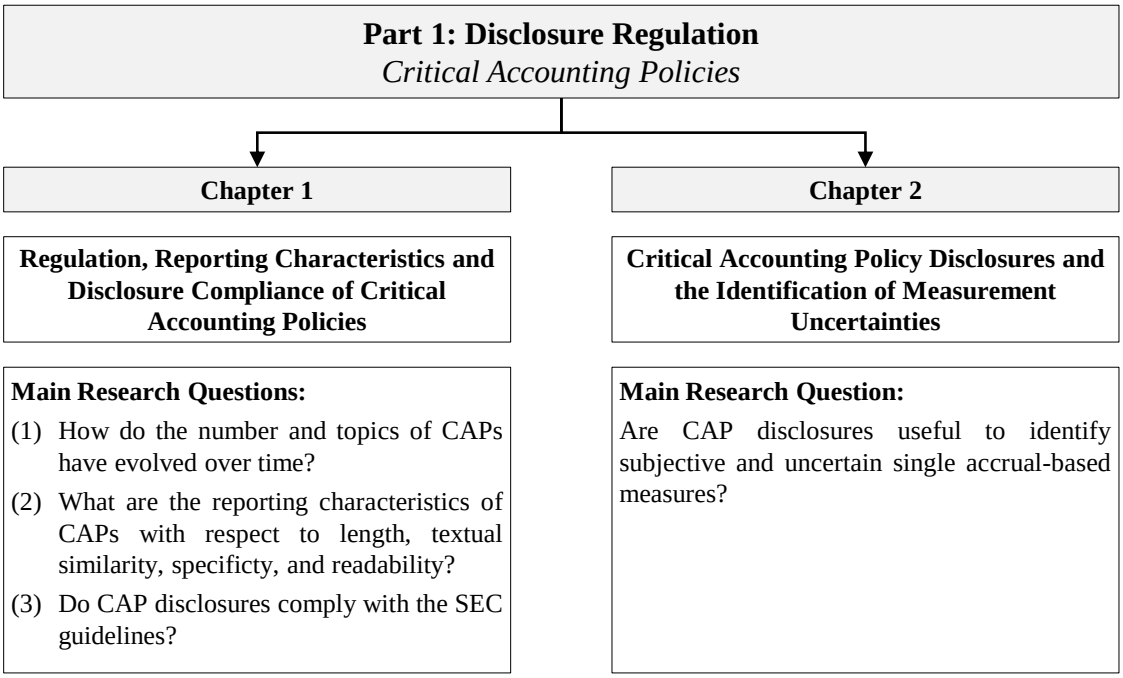
particular the issue of board diversity. These developments and discussions triggered my deep interest in disclosures and corporate governance. In this dissertation, I focus especially on two major topics that are imposed and discussed by standard setters, regulators, and the international press to improve the transparency and effectiveness of corporate disclosures and corporate governance.

The first part of this dissertation (Chapter 1 and 2) addresses one initiative of the Securities and Exchange Commission (SEC) in the early 2000s, which focuses especially on the interplay of firm-specific measurement uncertainties in the application of complex and highly uncertain accounting policies that are used in the preparation of financial statements. The main goal of the SEC is to improve the understanding of financial statement users about measurement uncertainties embedded in financial statements and, particularly, in accounting estimates. Since 2001, the SEC encourages firms to provide detailed information about their '*critical accounting policies*' (CAPs), which are those highly uncertain and complex accounting policies with a material impact on a firm's financial condition. Focusing on CAPs is of great interest because of two reasons. First, there is limited knowledge in this field of research as only a few studies have analyzed CAPs to this day. Second, there is no streamlined disclosure regulation within the International Financial Reporting Standards (IFRS). An article published in a German practitioners journal by Fülbier et al. (2017)[9] compares the CAP regulation with the disclosure requirements about estimation uncertainties according to International Accounting Standards (IAS) 1.125. We find that there are some commonalities between CAPs and estimation uncertainties. However, the SEC requires more detailed and profound information about measurement uncertainties. Interestingly, the IASB is currently discussing on how to improve disclosures about accounting policies, estimates, and estimation uncertainties (IASB 2019). Thus, providing further evidence about CAPs, which do not exist within the IFRS, might provide new interesting implications for standard setters that could enrich future discussions on accounting policy disclosures.

[9] The second author of Fülbier et al. (2017) is the same author of this dissertation.

Each chapter of the first part of this dissertation fills one gap in the literature of CAPs. While both chapters are based on hand-collected data and share a quantitative-empirical approach, each one offers distinct features in the data and research design. The first analysis is grounded on quantitative content analysis (Krippendorff 2019) and descriptive statistics (Smith 2017). The second chapter employs an empirical-archival analysis on multivariate grounds (e.g., Merchant 2010; Oler et al. 2010). Figure 1 depicts an overview of Chapter 1 and 2.

Figure 1
Overview Chapter 1 and Chapter 2



Chapter 1[10] provides the first comprehensive study about the regulatory framework, existing research, and implementation in practice of CAP disclosures. Focusing on a 16-year period from 2001 to 2016 and using hand-collected data about 35,686 CAPs, I provide initial evidence on the occurrence of highly uncertain accounting policies and estimates over time. Furthermore, my study is the first focusing on the (qualitative) reporting characteristics of CAP disclosures. Using tools from computational linguistics, I characterize CAP disclosures across a number of distinct dimensions that stand in line with the evolving textual analysis literature (e.g., Lang and Stice-Lawrence 2015) and are in the

[10] An earlier version of Chapter 1 has been presented at the 14th Workshop of Financial Reporting (EUFIN) in Stockholm. A paper version of Chapter 1 is available as Rupertus (2019).

interest of investors and regulators (e.g., Holtzmann 2007; SEC 2003a, 2016). In particular, I determine the length, textual similarity, specificity, and readability of CAP disclosures. There is a lack in prior research analyzing whether CAP disclosures currently comply with the SEC guidelines. Studying the content of CAP disclosures is interesting because to this day, the content of CAPs is still left to managerial discretion due to missing legal requirements. For this purpose, I conduct a content analysis of CAP sections of the largest 100 U.S. companies and expose how disclosures have been complied with the SEC guidelines in 2016.

My results provide several interesting new insights about CAP disclosures. First, firms provide six CAPs on average that mostly relate to deferred taxes, intangibles, property, plant and equipment, retirement benefits, revenue recognition and contingencies. Furthermore, accounting topics that are flagged as CAPs vary between firms and industries. While some CAPs are more common (e.g., retirement benefits, property, plant and equipment, and intangibles), others occur only in certain industries (e.g., warranties, financial instruments, and inventories). Nonetheless, CAPs do not vary significantly over time. Second, the average length of the complete section and for each CAP separately has constantly increased between 2001 and 2016. In accordance with prior literature, I argue that the prevention of litigation (Levine and Smith 2011), a higher precedence, or an intensive enforcement of the SEC (Cassell et al. 2013) might explain this development. The textual similarity of CAP disclosures is very high, indicating that the occurrence of uncertainties embedded in accounting estimates is relatively stable over time and that firms provide nearly identical disclosures each year. Moreover, CAP disclosures are largely unspecific (on average two specific words out of 100) and highly complex (i.e., difficult to read). Whereas prior studies find that textual similar disclosures have a positive effect on a firm's information environment (e.g., Peterson et al. 2015), it seems that specificity and readability only fulfil the requirements of the SEC in certain parts. Third, in comparison to the early results of prior literature (e.g., Hughes et al. 2009), I show that the content of CAP disclosures seem to have qualitatively improved in 2016, measured by the compliance with the subject's required by the SEC. In particular, I find that 98 percent of my sample firms include at least one CAP that comply with nearly 50 percent of the subjects required by the SEC. The majority of CAPs contain information with respect to the methodology, assumptions, as well as factors affecting the underlying assumptions and methodology.

Chapter 2[11] focus on the usefulness of CAP disclosures. This part is a joint project with Marcus Bravidor. We answer a call for additional research and extent the literature by examining whether CAPs fulfil their intended purpose and depict instances of measurement uncertainties embedded in individual financial statement positions. Existing literature finds that accruals are less useful in predicting future cash flows due to subjectivity in their estimation (e.g., Richardson et al. 2005; Sloan 1996). Following prior findings, we examine whether 'uncertain' (if the component is flagged as a CAP) accruals are less persistent with respect to future cash flows than accruals that are 'certain' (if the component is *not* flagged as a CAP). We argue further that the effect of measurement uncertainties on the persistence of uncertain accruals varies across firms and industries. Accordingly, we analyze whether the lower persistence of uncertain accruals depends on their importance and specificity for a given firm.

Using the disaggregated cash flow prediction model of Barth et al. (2001), we find that accruals flagged as CAPs are less persistent with respect to future cash flows. This finding is consistent with the SEC's intention that CAPs capture instances of greater measurement uncertainty embedded in the underlying accrual measurement. In an additional analysis, we provide initial evidence and find that uncertain accrual components are in fact not less useful in predicting future cash flows *per se*. It also depends on their importance and, to a certain extent, on the specificity for a given firm. As there is no evidence on how to identify subjective and uncertain accruals, we demonstrate that CAP disclosures provide such information. Our results are robust to alternative specifications of the model as well as to our measure of importance and specificity.

The second part of this dissertation (Chapter 3 and 4) focuses to the current debate about gender diversity on corporate boards. Board gender equality and the lack of female board representation has increasingly become the focus of international political and societal debates in recent years (e.g., Grosvold and Brammer 2011; Pande and Ford 2011; Singh et al. 2015; Terjesen et al. 2016). For instance, several European countries obligate a ratio of women on corporate boards to foster equal female participation in economic activity. In 2008, Norway adopted the first mandatory gender quota regulation. Similar require-

[11] An earlier version of Chapter 2 has been presented at the 40th European Accounting Association (EAA) Annual Congress in Milan. This version is accepted for presentation at the 41th EAA Annual Congress in Paphos and the 2019 Doctoral Colloquium in Larnaca. A paper version of Chapter 2 is available as Rupertus and Bravidor (2019).

ments have already been or will be adopted in Belgium, France, Germany and the Netherlands (Deloitte 2017; Terjesen et al. 2015). Gender quotas are stipulated on an international level in a wide range of voluntary corporate governance codes (Terjesen et al. 2015).[12] Various institutional investors demand a higher ratio of female board members and consider board diversity in their investment decisions (e.g., Cao and Donnelly 2010; Coffey and Fryxell 1991).[13] Despite these efforts, there still are considerable differences in terms of average female board membership (e.g., Loy and Rupertus 2018a; Terjesen and Singh 2008). Yet, academic research lacks studies that elaborate on why women in some countries are far more underrepresented in the boardroom. Furthermore, most of the above-mentioned initiatives are based on the view that board gender diversity could enhance corporate performance (e.g., Grosvold et al. 2016). Existing empirical research focuses on the association of female board representation with accounting performance and short-term market reactions. These studies neglect the long-term market performance attributable to board gender diversity. The following two chapters address these gaps in the literature. Whereas, both chapters are grounded on multivariate empirical-archival analysis, Chapter 3 uses data on the country-level, while Chapter 4 focuses on the firm-level perspective. Figure 2 depicts an overview of Chapter 3 and Chapter 4.

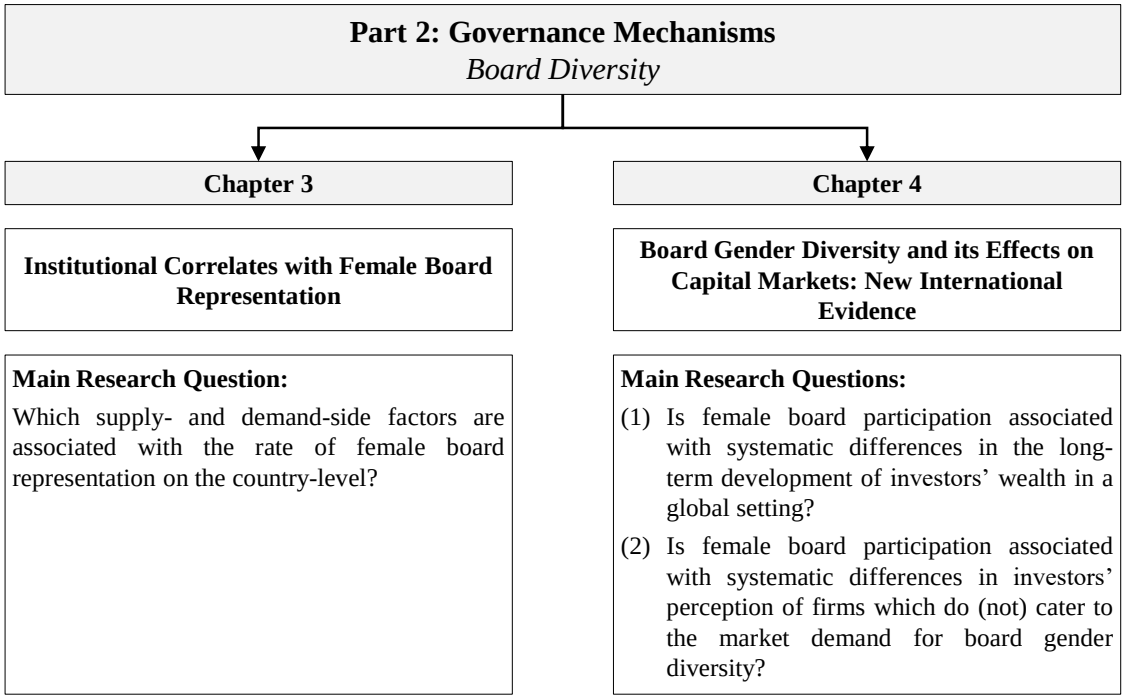
Chapter 3[14] is a joint project with Thomas R. Loy. We systematically and empirically disentangle institutional supply- and demand-side factors associated with global differences in female board representation. Our analysis is based on the *World Economic Forum's* Global Gender Gap Index (GGGI) that captures the overall level of societal gender equality across countries. The score captures 14 variables covering topics such as health and survival, political empowerment, economic participation, as well as opportunity. We classify these individual indicators as either supply or demand factors and extract both components using confirmatory principal component analysis (PCA).

[12] For an overview about mandatory gender quotas and voluntary corporate governance codes, cf. Deloitte (2017).

[13] For instance, the mutual fund 'Pax Global Women's Leadership Index' only invests exclusively in corporations that have established strict guidelines for female representation.

[14] Chapter 3 has been published as Loy and Rupertus (2018a).

Figure 2
Overview Chapter 3 and Chapter 4



Our result show that functioning outside investor protection and a societal climate of gender equality (GGGI) contributes to more women on boards, mainly by fostering the supply of suitable candidates. Furthermore, we provide initial evidence that longer-term supply-side factors need to complement short-term demand-side regulation. While gender quotas might regulate the demand for female upper echelons, they are ineffective to increase the supply of suitable candidates in the short run. Thereby, legislators might rather focus on supply-side measures, such as education, and giving women more opportunities to move into managerial and professional roles.

Chapter 4[15] focus on the association of female board participation and shareholders' wealth. This chapter is a joint project with Thomas R. Loy. We examine investors' perception and long-term effects of board gender diversity on firms' capital market performance in an international setting. Building upon role congruity theory, our study contributes to the social and economic debate about board gender diversity by analyzing how stakeholders (e.g., investors) perceive female board members in the long-term. Increased

[15] An earlier version of Chapter 4 has been presented at the 38th EAA Annual Congress in Maastricht and the 78th Annual Congress of the German Academic Association of Business Administration in Munich. A paper version is available as Loy and Rupertus (2018b) and is currently under review at *Business & Society* (3rd round).

board gender diversity demanded from a wide variety of stakeholders, such as institutional investors (Byoun et al. 2016; Coffey and Fryxell 1991), stock exchanges (Terjesen et al. 2016) as well as regulators (SEC 2009). Subsequently, catering theory argues that firms cater to investors' and other stakeholders' demands by appointing women to their boards (Ghosh et al. 2016). We evaluate further whether societal pressures result in firms fulfilling these expectations or punishments by investors if firms fail to 'comply'.

Our results indicate that female board representation neither improves nor reduces firms' long-term stock performance. Investors seem to perceive female and male board members as being equivalent in the long-term and do not base their investment decisions on directors' gender. Moreover, we fail to document significant reduced stock return for firms that deviate from the expected ratio of female board membership and vice versa. While female board appointments may have reported short-term effects (e.g., Kang et al. 2010; Lee and James 2007; Schmid and Urban 2016), it seems that the market corrects this (negative) mispricing over the long run.

CHAPTER 1:
REGULATION, REPORTING CHARACTERISTICS AND DISCLOSURE COMPLIANCE OF CRITICAL ACCOUNTING POLICIES

Abstract

Since 2001, U.S.-firms have been encouraged to disclose all highly uncertain accounting policies with a material impact on the presentation of the financial condition of the firm ('*critical accounting policies*', CAPs). Using a sample of Standard & Poor's (S&P) 500 firms between 2001 and 2016, I find that the number of CAPs has increased over time and that there is a huge variation in uncertain accounting topics across firms and industries. Moreover, I provide first evidence on (qualitative) reporting characteristics of CAPs. Specifically, CAP disclosures are very similar over time, remain largely unspecific, and are difficult to understand. Furthermore, the content improved qualitatively in 2016, as measured by the level of compliance with the subjects required by the Securities and Exchange Commission (SEC). Overall, the results expand the understanding about the occurrence, regulatory framework, and content of CAPs. In particular, it further allows indications about factors that determine a firm's decision to flag an accounting policy as critical, potential effects of CAPs on capital markets, and their usefulness for financial statement users.

A paper version of Chapter 1 is available as Rupertus (2019).

Acknowledgements: I gratefully acknowledge comments and suggestions by Marcus Bravidor, Rolf Uwe Fülbier, Thomas Loy, Klara Lösse, Christian Wittmann, delegates at the 2018 Workshop of Financial Reporting (EUFIN) in Stockholm, and seminar participants at the University of Bayreuth. Thanks to Brian Bloch for his editing of the English. I also thank Florian Federsel, Elisabeth Kuhn, Simon Lemnitzer, Marie-Thérèse Meyer, and Anna Mollat for their excellent research assistance.

1 Introduction

In the early 2000s, the Securities and Exchange Commission (SEC) started a series of initiatives to enhance the quality and transparency of corporate disclosures, in order to make capital markets more efficient (Hughes et al. 2009; SEC 2001). One initiative focused especially on the interplay of firm-specific uncertainties in the application of highly complex accounting policies and estimates that are used in determining financial statement positions. Since 2001, firms have been encouraged to disclose their '*critical accounting policies*' (CAPs), which are those policies requiring "management's most difficult, subjective, or complex judgements" (SEC 2001, p. 1) with a material impact on a firm's financial condition. The SEC suggested that disclosures about CAPs would increase investor understanding of a firm's financial condition, enabling more informed investment decisions (e.g., Hughes et al. 2009; SEC 2001). Guidance related to CAP disclosures is included in several SEC releases (Financial Reporting Release (FR) 60, the Proposed Rule and FR-72) from 2001 to 2003, requiring firms to provide detailed information about the uncertainties underlying their accounting estimates, as well as the effect on a firm's financial condition in the Management, Discussion & Analysis (MD&A) section of each 10-K (SEC 2001, 2002a, 2003a).

Prior empirical studies in this field have primarily analyzed the relationship between CAP disclosures and earning properties or economic outcomes (e.g., Glendening 2017; Levine and Smith 2011, cf. Chapter 2). However, to this day, there is only limited evidence on the implementation in practice of CAPs as well as on the information content provided in each CAP section. First, only a few studies analyze which accounting policies and estimates are flagged as CAPs and how related disclosures comply with SEC guidelines (e.g., Bauman and Shaw 2014; Hughes et al. 2009; Levine and Smith 2011). Second, most prior research focuses on single years immediately after the initial SEC releases and is based on single CAPs or on relatively small samples. To my surprise, previous studies have generally failed to analyze how CAP disclosures have evolved over time and how firms respond to SEC guidelines in later years. Furthermore, there is no prior research analyzing the qualitative information content of CAPs. Thus, current knowledge in this field is limited, so that in order to shed more light on this issue, I attempt to answer the following research questions: (1) How have the number and nature of accounting topics on CAPs

across firms and industries evolved over time, (2) what are the reporting characteristics of CAPs and (3) how do they comply with the SEC guidelines.

Focusing on my research questions, I provide several important contributions. First, whereas prior studies focus on single years or in particular on the years immediately after 2001, I focus on a 16-year period between 2001 and 2016. Therefore, I provide new holistic evidence about the occurrence of highly uncertain accounting topics as well as how they changed over time. Second, the SEC has repeatedly emphasized the importance of providing clear, insightful and understandable information. While quantitative information about CAP disclosures has been studied before, there is, to the best of my knowledge, no study analyzing the qualitative characteristics of CAPs. Focusing explicitly on the qualitative attributes of CAPs provides new evidence on the extent to which financial statement users might assess the information presented in each CAP section, as well as the impact of estimates, judgements, and uncertainties on a firm's financial status. Using tools from computational linguistics, I characterize CAP disclosures across a number of distinct dimensions that are in line with the evolving textual analysis literature (e.g., Lang and Stice-Lawrence 2015; Loughran and McDonald 2016). Specifically, I determine the length[16], textual similarity, specificity, as well as readability of CAP disclosures. Third, the SEC still emphasizes to include CAP disclosures in Regulation S-K and is attentive to any noncompliance with their releases (Cassell et al. 2013; Holtzmann 2007). Due to the missing legal requirements, FR-60, the Proposed Rule, as well as FR-72 still serve as major guidelines in the preparation of CAP disclosures and thus, the content of each CAP section is still left to the discretion of each firm. There is a lack of research on whether current CAP disclosures comply with SEC guidelines. Studying the content of CAP disclosures provide new evidence on disclosure compliance with the information provided in each CAP section, as well as on whether the SEC should strive to provide further releases to enhance the quality of CAP disclosures.

First, my results show that the number of highly uncertain accounting topics increased significantly in the first years after the introduction of CAPs and remain similarly high in subsequent years. On average, firms provide about six CAPs that mostly relate to deferred taxes, intangibles, property, plant and equipment, retirement benefits, revenue recognition, and contingencies. Whereas the identified accounting topics flagged as CAPs do not

[16] Although the average length of CAPs is a quantitative measure, I present related results with the other qualitative characteristics.

change significantly over time, there are huge discrepancies at the industry level. Thus, industry seems to be a significant explanatory factor with respect to some CAPs. Second, I find that the average length of the complete section, as well as for each CAP separately, has increased constantly between 2001 and 2016. Firms gradually disclose more information about their CAPs over time. The prevention of litigation (e.g., Levine and Smith 2011), a higher priority or intensive enforcement by the SEC (e.g., Cassell et al. 2013) might be reasons for this development. Furthermore, the textual similarity of CAP disclosures is very high over time, indicating that the occurrence of uncertainties embedded in accounting estimates is relatively stable and that firms provide almost identical disclosures each year. Moreover, CAP disclosures are largely unspecific (on average only two specific words out of 100) and are extremely complex in terms of readability (i.e., difficult to understand). Whereas prior studies find a positive association between textually similar disclosures and a firm's information environment (e.g., Peterson et al. 2015), it seems that specificity and readability only fulfil the requirements of the SEC to a certain extent. However, further research is needed to empirically analyze the association between the aforementioned characteristics and firm fundamentals, analyst data, or economic outcomes, in order to make a more comprehensive statement about whether CAPs provide useful information to outsiders. Third, I conclude that the content of CAP disclosures improved qualitatively in 2016, compared to the prior literature, measured by compliance with the subjects required by the SEC. Most CAP disclosures conform to the requirements included in SEC releases. I further find that 98 percent of my sample firms include at least one CAP that complies with almost 50 percent of the subjects required by the SEC. Moreover, the majority of CAPs contain general information with respect to the methodology, assumptions, as well as factors affecting the underlying assumptions and methodology.

I structure the remainder of the paper as follows. First, I describe the regulatory framework, discuss prior literature, and derive my research questions. Second, I describe my sample selection and methodological approaches. Third, I present and discuss my results. Fourth, I derive practical implications as well as fruitful avenues for future research. The final section concludes.

2 Regulatory Framework about Critical Accounting Policies

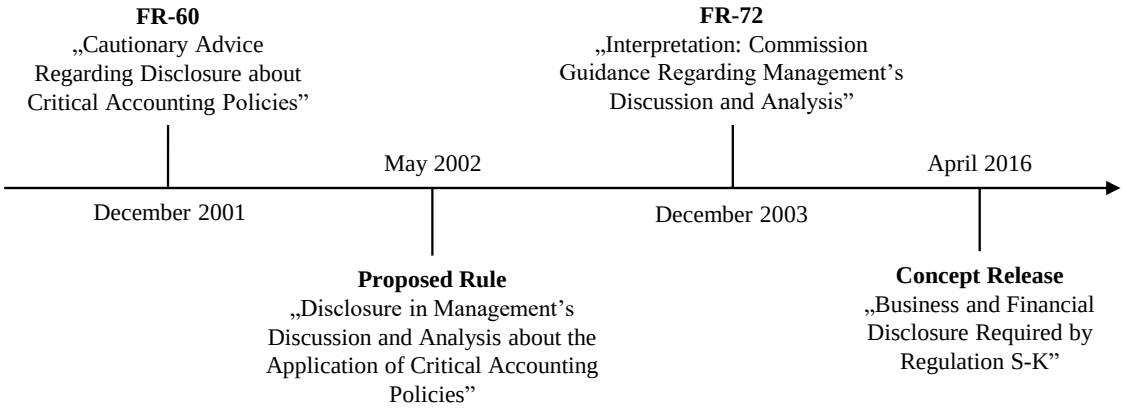
In response to a call for more transparent information regarding accounting methods, assumptions, and estimations, the SEC announced new disclosure requirements in the early 2000s to enhance investor understanding of judgements, assumptions, and uncertainties affecting the application of accounting policies and estimates with a material impact on a firm's financial condition (SEC 2001). Specifically, the SEC focused on additional disclosures other than those required by the Financial Accounting Standards Board (FASB) in Accounting Standards Codification (ASC) 235-10-50[17] and ASC 275-10-50[18] and encouraged firms to disclose those accounting policies and estimates requiring "management's most difficult, subjective, or complex judgements, often as a result of the need to make estimates about the effect of matters that are inherently uncertain" (SEC 2001, p. 1) (CAPs). The main goal of the new regulation is to provide greater insights into the interplay of highly uncertain accounting estimates, operating performance, and a firm's financial condition "through management's eyes" (SEC 2002a, p. 9). Thus, CAP disclosures are included in the MD&A section, rather than in the notes to the financial statements. Figure 3 presents a timeline of SEC releases that focus on CAP disclosures.

The initial guidance of this new disclosure regulation is included in FR-60, *Cautionary Advice Regarding Disclosure about Critical Accounting Policies*. FR-60 was released in December 2001 and encourages firms to include a full explanation of the judgements and uncertainties affecting the application of accounting principles, as well as the likelihood that materially distinct amounts are reported under different conditions or by using different assumptions (SEC 2001). Nevertheless, FR-60 does not contain specific guidance with respect to the implementation of CAP disclosures and thus, the SEC announced that it would be releasing further clarifications (Hughes et al. 2009; SEC 2001, 2002b).

[17] ASC 235-10-50 requires firms to identify and describe all significant accounting policies, methods, and judgements that are required in the valuation of financial statement positions and have a material effect on a firm's financial situation (Flood 2018). The accounting policy section should also include the selection from existing acceptable alternatives, industry specific methods in which the firm operates or unique and unusual applications of accounting principles (ASC 235-10-50-3 (a) – (c)).

[18] ASC 275-10-50 provides guidelines that should help outsiders to identify risks and uncertainties in the preparation of financial statements (Flood 2018). On the one hand, firms should provide an explanation of the preparation of financial statements and information required about the use of estimates (ASC 275-10-50-1 (b); ASC 275-10-50-4; Flood 2018). On the other hand, ASC 275-10-50-1 (c) requires a discussion of estimates when it is reasonably possible that they will change soon and would have a material effect on the financial condition (ASC 275-10-50-6; ASC 275-10-50-8). Additionally, a company's disclosure shall encompass the nature of the uncertainties as well as an indication that it is reasonably likely that a change in the estimate will occur (ASC 275-10-50-9).

Figure 3
Timeline of CAP Regulation



One year later, in May 2002, the SEC issued a Proposed Rule entitled "*Disclosure in Management's Discussion and Analysis about the Application of Critical Accounting Policies*" containing detailed guidelines on quantitative as well as qualitative disclosures about a firm's CAPs (SEC 2002a). The rules explicitly distinguish between '*critical accounting estimates*' (CAE) and CAPs. Whereas the former are defined as judgmental and subjective estimates involved in the application of (critical) accounting policies with a material impact on a firm's financial condition, the latter are those accounting policies which require management's most difficult, subjective, and complex judgements (SEC 2002a). However, empirical evidence shows that companies still do not differentiate adequately between both terms (e.g., Fülbier et al. 2017). Thus, the terms CAEs and CAPs are used interchangeably within this study and I will refer mainly to CAPs. The primary goal of the Proposed Rule is to increase the transparency of CAP disclosures so that investors would gain a greater understanding about highly subjective and complex accounting policies and estimates. Consequently, financial statement users might better assess the quality as well as potential variability of current and future earnings (SEC 2002a). According to the Proposed Rule, each section should include disclosures about the nature, methodology, assumptions, and significance of each CAP. Firms should also disclose a qualitative and quantitative analysis about the sensitivity of each estimate and how earnings would be affected by changing an uncertain estimate, if material. Furthermore, each firm should include an explanation about whether the selection and application of CAPs were discussed with the audit committee as well as a discussion on a segment basis

(Holtzmann 2007; SEC 2002a). Due to extensive criticism for obscuring rather than revealing information to investors in their decision-making (e.g., Sullivan and Cromwell 2002), the Proposed Rule was not adopted within Item 303 of Regulation S-K (Bauman and Shaw 2014).

In 2003, the Division of Corporate Finance reviewed the 10-Ks of Fortune 500 companies and focused primarily on disclosures according to the recommendations included in FR-60. As an overall result, the SEC noted that a substantial number of companies did not provide any CAP disclosures. However, in case of CAP disclosures, they were not adequately congruent with the SEC guidance (SEC 2003b). To provide further guidance, the SEC released FR-72, *Interpretation: Commission Guidance Regarding Management's Discussion and Analysis of Financial Condition and Results of Operations*, in December 2003. The commission emphasized that the description of CAPs should supplement, and therefore not duplicate the accounting policy section that is already disclosed within the notes to the financial statements. While the notes about accounting policies generally describe the methods used to apply accounting principles (ASC 235-10-50), each CAP section within the MD&A should provide an analysis of the company's uncertainties involved in applying their accounting policies and estimates. Furthermore, a company should provide a sensitivity analysis and a discussion regarding the accuracy of past and future estimates. Firms further have to add how they arrived at the estimate (SEC 2003a). However, most disclosure subjects in FR-72 reproduce the contents of the Proposed Rule (Hughes et al. 2009).

In 2016, as a reaction to the evolving criticism regarding the usefulness of overall corporate disclosures, the SEC published a Concept Release, *Business and Financial Disclosure Required by Regulation S-K* (S7-06-16), to assess whether Regulation S-K still includes guidelines to provide decision useful information to investors (SEC 2016). One part of the Concept Release focuses on CAP disclosures. Based on eight questions, the Commission strived to receive feedback from users and firms on whether they should revise Item 303 to mandate CAP disclosures and how to make them more informative for investors.[19] Overall feedback in response to the comment letters has been mixed. While most respondents support the idea of revising Item 303 of Regulation S-K to mandate CAP disclosures (e.g., California Public Employees' Retirement System 2016; Center for

[19] To date the commission has received 376 comment letters, of which only 32 address CAP disclosures.

Audit Quality 2016; PwC 2016), there are others stating that there is no need to include a CAP section in the MD&A (e.g., Fenwick & West LLP 2016; Chevron Corporation 2016). Table 1 presents an overview of the recommended disclosures about CAPs (SEC 2001, 2002a, 2016).

Table 1

Content of CAP Disclosures according to FR-60, the Proposed Rule and FR-72

Subject
Description about the nature of the estimate, how firms arrived at the estimate, the methodology, and material assumptions that are highly uncertain in the application of the estimate.
Explanation of all trends, circumstances, and factors that materially affect the application of the methodology and assumptions.
Quantitative and qualitative information about the impact of the estimate on the company's financial condition and results of operations.
Identification of the financial statement line items that are affected by the CAP.
Quantitative and qualitative information about the accuracy of the estimate in the past.
Quantitative and qualitative information about material changes made to the CAP in the past three years.
Quantitative discussion about the sensitivity of the estimate with respect to the overall financial performance.
Statement of whether or not the selection and development of the CAP was discussed with the audit committee.
A discussion of the accounting estimates on a segment basis.

There have been several SEC releases emphasizing the importance of communicating highly uncertain accounting estimates and assumptions, as well as providing companies with guidelines for improving their disclosures, and this continues to this day. Despite all regulatory efforts, no final rule has in fact been published and it remains unclear whether the Commission will revise Item 303 of Regulation S-K to mandate CAP disclosures in the near future. However, overall opinion from the comment letters received from the Concept Release is that CAP disclosures are helpful and that the SEC should incorporate principal-based requirements to enhance investor understanding of the measurement process of highly complex accounting estimates and policies and their impact on financial statements (e.g., California Public Employees' Retirement System 2016; Center for Audit Quality 2016; PwC 2016).

3 Prior Literature

This section provides an overview of research on CAPs. For this purpose, I collect and review eight published articles and three working papers devoted mainly to CAP disclosures. All identified studies were published between 2004 and 2019. Table 2 depicts the reviewed literature.

To date, research on CAP disclosures has focused mainly on the following subjects: Descriptive statistics on the number of CAPs and related accounting topics, quality of CAP disclosures and its association with earning properties and economic outcomes, the content, as well as determinants of CAP disclosures. Studies presenting descriptive findings about the accounting topics find that deferred income taxes, revenue recognition, pensions, property, plant and equipment, financial instruments, as well as impairments are the most frequent CAPs and that most firms disclose around five to six CAPs (Cho et al. 2005; Fülbier et al. 2017; Holtzmann 2007; Levine and Smith 2011; Paprocki and Stone 2004). Furthermore, only two studies analyze the disclosure quality of CAPs and its association with economic outcomes or earnings properties. Using a self-constructed disclosure index based on the SEC guidelines, Paprocki and Stone (2004) conclude that higher disclosure quality is associated with an improved information environment. Cho et al. (2005) show that the quality of CAP disclosures varies both across and within industries and is positively associated with accrual quality. Both studies provide some evidence that CAP disclosures contain information about the underlying accrual positions, and might be useful for reducing information asymmetries. The content of CAP disclosures has mostly been measured as compliance with the evolving SEC guidelines (e.g., Hughes et al. 2009; O'Shaughnessy and Rasthy 2005) or single subjects required by the SEC (such as a quantitative discussions about the sensitivity of the underlying estimates) (e.g., Bauman and Shaw 2014; Glendening 2017). While the compliance with the SEC guidelines improved between 2001 and 2003, the aforementioned studies conclude that various subjects mentioned in the releases remain underdisclosed. Thus, there might be room for improvement.

Table 2
Research on CAP Disclosures

Study	Subject	Sample	Main Variables	Main Results
Paprocki and Stone (2004)	Disclosure quality and economic outcomes	293 S&P 500 firms (2001-2003)	*Disclosure Quality*: Self-constructed disclosure index (0-4) based on SEC guidance *Information asymmetry*: (1) Altman's Z bankruptcy prediction score (2) Book-to-market ratio (3) Proportion of outstanding debt	On average, firms disclose three to five CAPs and the most frequently disclosed ones relate to impairment of long-lived assets, pensions, deferred income taxes, allowance for doubtful accounts, and revenue recognition. Disclosure quality of CAPs increased from 2001 to 2003, and a higher disclosure quality is associated with an improved information environment (i.e., lower information asymmetry).
Cho et al. (2005)	Disclosure quality and earnings properties	500 U.S. firms (2002)	*Disclosure Quality*: Self-constructed disclosure index (1-5) based on SEC guidance *Accrual Quality*: (1) Discretionary accruals (Jones 1991) (2) Standard deviation of the residuals from the Dechow and Dichev (2002) model	Most frequently disclosed CAPs are revenue recognition, accounting for goodwill, pension accounting, property, plant and equipment, and environmental liabilities. Pension-related CAPs have the highest disclosure quality, followed by investments in financial securities, revenue recognition, and accounting for derivatives. Quality of CAP disclosures varies across firms and industries and is positively associated with accrual quality.
O'Shaughnesy et al. (2005)	Disclosure content	10 NASDAQ firms (2003)	Checklist with the SEC guidance	CAP disclosures are more robust than related disclosures in the notes to the financial statements. Most CAPs only contain a short description of the underlying GAAP rules. Companies are complying with the statutory disclosure requirements but sometimes fail to disclose the level of judgement in exercising the accounting estimate.

Table 2 - continued
Research on CAP Disclosures

Study	Subject	Sample	Main Variables	Main Results
Holtzmann (2007)	Disclosure content and accounting topics	100 Fortune 500 firms (2005-2006)	Number of CAPs and accounting topics	Most frequent disclosed CAPs are impairment for intangibles, pensions, and income taxes, followed by contingencies, revenue recognition, and financial instruments. Moreover, there are several industry-specific CAPs, e.g., claims, oil and gas accounting, and capitalization of entertainment assets.
Hughes et al. (2009)	Disclosure content and disclosure quality	112 Mid-Cap 400 index firms (2001 and 2003)	Checklist with the SEC guidance *Disclosure Quality*: Number of sentences	Some firms in 2001 and all firms in 2003 include CAP disclosures. Most firms comply with at least some of the subjects required by the SEC. The disclosure content and quality increased from 2001 to 2003, but even two years after FR-60 and FR-72, several subject matters remain underdisclosed. The number of content-specific sentences, the number of subjects required by the SEC, as well as the length of the disclosures, increased from 2001 to 2003.
Levine and Smith (2011)	Disclosure content, accounting topics, determinants, earning properties, and market reactions	4,397 U.S. firms in 2003	Probability of litigation, firm's average financing activity, historical mean and variance of the CAP, unexpected CAP disclosures	The most frequent CAPs are marketable securities, asset impairment, and revenue recognition. Furthermore, firms with ex-ante higher litigation risk are more likely to provide CAP disclosures, indicating that firms use this disclosure practice to reduce their exposure to lawsuits. The decision to provide CAP disclosures is associated with the current and future variance and magnitude of the financial statement accounting. The authors further conclude that firms with fewer (more) CAPs than expected report more (less) reliable earnings. In addition, when the number of CAPs is greater (lower) than expected, market returns are likely to reverse around the filing date (are in line with the earnings announcement returns).

Table 2 - continued
Research on CAP Disclosures

Study	Subject	Sample	Main Variables	Main Results
Glendening (2012)	Value relevance and earning properties	339 S&P 500 firms (2004-2009)	Market value of equity, earning persistence, accrual persistence	CAE disclosures explain cross-sectional variation in the value relevance of the underlying financial statement positions. Investors perceive financial statements that are flagged as critical as less reliable (but only on the post-disclosure period). Earnings are less useful in predicting future cash flows when a CAE disclosure is present. Accrual distortion is greater when a highly material working capital estimate (CAE) exists.
Bauman and Shaw (2014)	Disclosure content, determinants	147 U.S. firms with large pension plans (2010)	Quantitative pension CAE disclosures, disclosure of the sensitivity of pension expense	Only 60 percent of the firms provide a quantitative sensitivity analysis of their pension-related CAEs. The propensity to provide a quantitative sensitivity analysis, in combination with a pension discount rate and expected asset return, is positively associated to firm size, BIG4 auditor, and the variability of pension plan funded status.
Fülbier et al. (2017)	Disclosure content	30 U.S. and 30 German firms (2015)	Number of CAPs and accounting topics	CAP disclosures require more profound information about highly uncertain accounting topics than related disclosure requirements according to International Accounting Standards (IAS) 1.125. German firms provide slightly more accounting topics that are affected by estimation uncertainties. The most frequent uncertain accounting topics are deferred taxes, revenue recognition, and impairments.
Glendening (2017)	Earning properties	339 S&P 500 firms (2004-2009)	Earning persistence	Earnings are less persistent with respect to future cash flows in the presence of a quantitative CAE disclosure.
Glendening et al. (2019)	Determinants	339 S&P 500 firms (2004-2009)	Management's reporting incentive, audit committee expertise, accounting misstatements	The decision to provide a quantitative sensitivity disclosure of CAEs reflects strategic preferences of those that are responsible for financial reporting. The disclosure of a quantitative sensitivity analysis is negatively associated with incentives to misreport and positively associated with audit committee accounting expertise. Accounting and Auditing Enforcement Releases (AAERs), misstatements, and earnings surprises are less likely after an initial quantitative CAE discussion.

As CAP disclosures vary between industries and firms, one strand of prior studies focused especially on the determinants of CAPs. Whereas Levine and Smith (2011) argue that firms use CAPs to mitigate litigation, Glendening et al. (2019) find that a firm's decision to provide quantitative sensitivity disclosures reflects strategic preferences of those that are responsible for financial reporting and is negatively (positively) associated with incentives to misreport (with audit committee accounting expertise). Another strand of studies focuses on the association of CAP disclosures with earning properties, such as earnings and accruals persistence, or how market participants perceive financial statement positions that are flagged as CAPs. For instance, Glendening (2017) shows that earnings are less useful in predicting future cash flows, if a firm provides quantitative sensitivity disclosures about highly uncertain accounting policies and estimates. Glendening (2012) provides evidence that investors perceive financial statement positions that have been flagged as CAPs as less reliable. Levine and Smith (2011) conclude that market returns are more likely to reverse and firms have less reliable earnings if they provide more CAPs than expected. By comparing CAP disclosures with related disclosures in the notes to the financial statements, O'Shaughnessy and Rasthy (2005) state that CAP disclosures are far more robust than related footnote disclosures. Moreover, Fülbier et al. (2017) find that the SEC requires more profound information about a firm's CAPs than the International Accounting Standards Board (IASB) about estimation uncertainties according to International Accounting Standard (IAS) 1.125.

To sum up, whereas one strand of prior studies focuses mainly on the relationship to earning properties, economic outcomes, and market perceptions, only some studies analyze how CAP disclosures are compliant with the SEC guidance and how the implementation in practice of CAPs have evolved over time. As can be seen in Table 2, the majority of prior studies focus on single years or on the years immediately after FR-60, the Proposed Rule, and FR-72. Thus, there is little evidence on the number and most frequent accounting topics across firms and industries, as well as on the compliance of CAP disclosures with the SEC guidelines, especially in later years and over time. To the best of my knowledge, a comprehensive empirical study analyzing CAP disclosures over a longer period has not yet been conducted.

4 Research Questions

The disclosure of CAPs has received great attention in the literature as well as from standard setters and practitioners (e.g., Bauman and Shaw 2014; Billings 2011; Glendening 2012; Glendening 2017; Henry and Holtzmann 2006; Levine and Smith 2011; Pitt 2002). Having said that, the majority of previous studies analyzing the implementation of CAP disclosures focus on single years immediately after the initial SEC releases or base their studies on relatively small samples. Due to all this, current knowledge on how firms respond to the SEC guidelines and communicate their CAPs over time is almost non-existent. This is significantly interesting, because, despite the fact that there have been no further releases after FR-72, the SEC still emphasizes the importance of providing informative CAP disclosures (SEC 2007a, 2016). Moreover, recent studies reveal that there is increasing complexity arising from the market, business, and accounting environment (e.g., Fülbier and Kuschel 2012; Kuschel 2015; Ojala et al. 2011) that might be reflected in a firm's CAPs. So far, there is no study analyzing how the implementation of CAPs has evolved and how firms communicate measurement uncertainties embedded in accounting numbers over a longer period. Explicitly taking a longer-term perspective may provide useful new insights into the distribution and variation of highly uncertain accounting topics across firms and industries. Based on this lack of research, I state my first research question as follows:

RQ 1: *How do the number and topics of CAPs have evolved over time?*

Second, CAP disclosures are narrative and included in the MD&A section of a firm's 10-K. The SEC and practitioners repeatedly emphasize the importance of providing clear, insightful, specific, and understandable information in plain English (e.g., Herdman 2002; SEC 2003a, 2016). Furthermore, firms should use simple sentences, avoid complex words, and use the active rather than the passive voice (Holtzmann 2007). However, analyzing to what extent CAP disclosures fulfil these requirements has so far received no attention in published research. I argue that the aforementioned attributes are important characteristics of CAP disclosures for two reasons: First, they show the extent to which financial statement users might assess the information presented in each CAP section. Second, they provide greater insights into the quality of related disclosures. To shed light on this issue, I use tools from computational linguistics to characterize CAP disclosures

across a number of distinct dimensions that are in conformity with the evolving textual analysis literature and are in the interest of investors and regulators in terms of affecting the information content (e.g., Lang and Stice-Lawrence 2015; Loughran and McDonald 2016). Specifically, I determine the length, textual similarity, specificity, as well as the readability of CAP disclosures. This procedure allows me to approximate the information content of CAPs across firms, industries and time. Thus, my second research question is as follows:

RQ 2: *What are the reporting characteristics of CAPs with respect to length, textual similarity, specificity and readability?*

Third, as shown in the previous section, only a handful of studies focus on the compliance of CAP disclosures with the SEC requirements. In this context, previous studies focus solely on the years immediately after FR-60 (Hughes et al. 2009) or on single accounting topics, such as retirement benefits (Bauman and Shaw 2014). Nonetheless, the SEC still emphasizes including CAP disclosures into Regulation S-K and is attentive to any non-compliance with their releases (Cassell et al. 2013; Holtzmann 2007; Glendening 2017). To the best of my knowledge, there is no evidence on how current CAP disclosures comply with the SEC guidelines and how the content varies between distinct topics. Thus, my third research question is as follows:

RQ 3: *Do CAP disclosures comply with the SEC guidelines?*

5 Sample, Data and Methodology

5.1 Sample Selection

To answer my research questions, I initially consider all firms of the Standard & Poor's (S&P) 500 as of 31.12.2016. The S&P 500 includes the largest listed companies in the U.S. with the highest market capitalization. This may provide useful new insights into the occurrence of highly uncertain accounting estimates and accounting positions of the largest and most important companies in the U.S. Moreover, S&P 500 firms may more often be the focal point of the SEC and thus, should strive to be compliant with their guidance regarding CAP disclosures. First, I start my sample selection by downloading all 10-Ks

of each firm of the initial sample from the Electronic Data Gathering, Analysis, and Retrieval (EDGAR) platform between 2001 and 2016. I rely on this period, because the SEC's first announcement of CAP disclosures was in 2001. Second, I limit my sample to firms with complete time series data. The final sample for my first and second research question consists of 402 firms and 6,432 firm-year observations. Furthermore, I consider the largest 100 companies in 2016 for my third research question. Table 3 depicts the sample selection procedure.

Table 3
Sample Selection

		No. of Observations
	Total number of observations with available 10-Ks between 2001 and 2016	7,303
-	Observations without complete time-series data	871
=	Finale Sample Research Question 1 and 2	6,432
=	Finale Sample Research Question 3	100

5.2 Methodology

5.2.1 Research Question 1

To answer my first research question, I extract the CAP section from each 10-K using Python.[20] Afterwards, I hand-collect each CAP heading and use a keyword-based coding system to assign each CAP to a single accounting position. My coding system consists of 30 accounting positions and is based on the FASB Taxonomy (FASB 2019) and the study of Levine and Smith (2011). In order to assign each CAP to an accounting topic, I select various keywords that pick up a related policy disclosure. To determine these keywords, I manually code CAPs from 100 firms between 2001 and 2016. I then code all remaining CAPs based on my defined keywords. This procedure enables a replicable coding approach. I present my coding system in Appendix A. Subsequently, I determine the average number of CAPs, as well as related accounting topics for each firm separately from 2001 to 2016. I present the most common CAPs for the whole sample and for each industry separately.

[20] Specifically, I use regular expressions to identify the CAP section within the MD&A. See Hering (2016) for a detailed explanation of how to retrieve textual information from 10-Ks. Because the raw text is in Hypertext Markup Language (HTML) format, I use regular expressions to parse the CAP section following the procedure of Loughran and McDonald (2019a).

5.2.2 Research Question 2

With respect to my second research question, I use four textual attributes in line with the evolving textual analysis literature (Hope et al. 2016; Lang and Stice-Lawrence 2015; Loughran and McDonald 2016) to characterize CAP disclosures. I first consider the length, measured by the number of characters in each CAP section. To control for the number of highly uncertain accounting topics, I determine the average length per CAP, measured as the length of each section divided by the total number of CAPs.

Second, textual similarity was introduced into the accounting literature by Brown and Tucker (2011). I measure textual similarity as the cosine similarity of a firm's CAP disclosures from year to year. Cosine similarity is based on the vector space model (VSM) which reflects the degree of similarity between two strings (Salton et al. 1975). The cosine similarity measures the cosine of the angle between two vectors that include unique words of a text after removing stop words and stemming remaining words (e.g., Peterson et al. 2015).[21] The cosine similarity (Equation 1) is defined as:

$$\cos\theta = \frac{V_1 \times V_2}{V_1 \times V_2} \tag{1}$$

V_1 and V_2 represent the vectors of two documents and V_i the vector norm $\overline{V_i \times V_i}$. The measure identifies similar documents by comparing the relative word frequencies across two documents and can have values between zero and one. If two documents are identical, the cosine similarity is one; if there are no overlapping words, the score is zero. This measure is widely used in accounting research to estimate the similarity of textual financial information (e.g., Bozanic and Thevenot 2015; Brown and Tucker 2011; Hoberg and Phillips 2010; Peterson et al. 2015).

Third, Hope et al. (2016) introduce a specificity score that enables researchers to assess the level of specificity of qualitative information. This measure captures another dimension of quality. Specificity is defined as "the number of specific words or phrases conveying specific information relevant to the disclosing firm, divided by the number of total

[21] Stop words include common words such as 'an', 'become' or 'among' that have no content. I use the Loughran and McDonald stop word list that is based on the Python's Natural Language Toolkit (NLTK) (Loughran and McDonald 2019b). The list is accessible on https://sraf.nd.edu/textual-analysis/resources/#StopWords (22.03.2019). Stemming remove suffices from words to obtain the 'stem' of the words.

words" (Hope et al. 2016, p. 1013). I follow Hope et al. (2016) and determine the specificity of each CAP disclosure by using the Named Entity Recognition (NER) technique and specifically the SpaCy toolkit.[22] NER is based on a natural language processing technology that allows finding and classifying elements of a text into predefined categories (Hope et al. 2016). The specific entity names of the SpaCy tool belong to 17 categories such as people, company names, percentages, monetary values, measurements, and dates. The higher the specificity score, the more specific the text. For instance, a score of 0.05 indicates that 5 out of 100 words are specific.

Fourth, financial reports should be understandable and written in plain English (e.g., SEC 1998, 2007b). Financial statements users should also be able to easily process the information they need to make their decisions (e.g., Li 2008; SEC 2000). Thus, processing costs might be reduced when disclosures are more informative and contain clear messages (Bloomfield 2002). Consistent with the literature as well as SEC guidance, I measure the readability of CAP sections regarding the extent to which they are clear and understandable. Readability is measured by using the Gunning Fog Index (thereafter named as the Fog index) (Gunning 1952). Up to the present, several prior studies use the Fog index to measure the text complexity and readability (e.g., Lawrence 2013; Li 2008; Miller 2010). The Fox index is based on the length of sentences as well as the proportion of complex words, where the latter refer to words with more than three syllables (e.g., Lang and Stice-Lawrence 2015). More specifically, the Fog Index (Equation 2) is defined as:

$$\text{Fog Index} = \left(\frac{\text{Number of Words}}{\text{Number of Sentences}} + \frac{\text{Number of Complex Words}}{\text{Number of Words}}\right) \times 0.4 \quad (2)$$

The higher the Fog index, the more complex (i.e., less readable) a text. A Fog Index greater than 18 implies that a text is unreadable and a score of 14-18 (less than 14) means that the text is difficult (easy) to read (Li 2008).[23]

[22] In comparison, Hope et al. (2016) use the Stanford NER tool that can extract seven specific entity names including names of people, locations, organizations, percentages, money values, times, and dates (The Stanford Natural Language Processing Group 2019). However, I use the SpaCy toolkit, because it provides a greater number of categories (ExplosionAI 2019). In an additional analysis, I also use the Sanford NER tool to measure the specificity of CAP disclosures. My main inferences remain qualitatively unchanged.

[23] According to Li (2008), a Fog Index of 12-14 implies that the text is optimally readable; 10-12 acceptable; and 8-10 childishly simple. However, the Fog index comes along with several limitations (e.g., Loughran and McDonald 2014). Amongst others, it considers neither the technical language of 10-Ks,

5.2.3 Research Question 3

I use content analysis to collect the data for my third research question. Content analysis is commonly used in prior studies to analyze how firms respond to new disclosure requirements as well as how the disclosure behavior has changed over time (e.g., Hughes et al. 2009; Marquardt and Wiedman 2007; Roulstone 1999). The aim of the technique is to describe, abstract, simplify, and structure the content of text to allow researchers to make specific inferences (e.g., Berelson 1952; Krippendorff 1978, 2019; Neuendorf 2017). To go beyond the data and abstract from the original complexity, text of the same or similar meaning are assigned into defined categories (e.g., Elo and Kyngäs 2008; Klein and Fülbier 2018; Krippendorff 2019). There are two ways to perform data categorization. First, the inductive approach enables researchers to process codes by identifying, refining, and validating categories directly from the text. Second, there is the deductive approach where categories are deduced *ex ante* from prior evidence, theoretical considerations, or regulatory frameworks (e.g., Elo and Kyngäs 2008; Mayring 2010). Because my third research question focuses on the compliance of CAP disclosures with SEC guidelines, I consider deductive content analysis and derive the categories based on the subjects mentioned in FR-60, the Proposed Rule, FR-72, the Concept Release, and the study of Hughes et al. (2009). To answer my third research question, I focus on CAP disclosures of the largest 100 firms of my sample in 2016. This enables me to analyze to what extent firms are compliant with the current guidelines. My sample size is in line with prior literature analyzing disclosure behavior with respect to SEC releases (e.g., Herrmann and Thomas 2000; Hughes et al. 2009; Roulstone 1999). To obtain the relevant data, I read all CAP sections and assigned the sentences to my predefined subject categories. After coding about 30 percent of the CAP sections, I recoded for intracoder-realiability-reasons the same data, and then, I resolved classification differences and coded the other sections. Afterwards, I went through all assignments again to ensure consistency.

nor the academic background of the reader. Furthermore, words like ‘business’ or ‘corporations’ are classified as complex words, yet are easy to understand for investors.

6 Results

6.1 Number and Topics of CAPs

Research question 1 focuses on the number as well as distinct accounting topics that are flagged as CAPs over time. First, I present the number of firms with and without CAP disclosures by year (Panel A) and industry (Panel B) in Table 4.[24] Table 5 presents descriptive statistics about the total number of CAPs, as well as the average number of CAPs per year (Panel A) and industry (Panel B).

Table 4
Sample Distribution

Panel A: By Year					
	All firms	Firms with CAP Disclosures	%	Firms without CAP Disclosures	%
2001	402	284	70.65	118	29.35
2002	402	395	98.26	7	1.74
2003	402	401	99.75	1	0.25
2004	402	402	100.00	0	0.00
2005 – 2016	402	402	100.00	0	0.00
Sum	6,432	6,306		126	
Panel B: By Industry					
NoDur	384	377	98.18	7	1.82
Durbl	128	125	97.66	3	2.34
Manuf	522	515	97.54	13	2.46
Enrgy	352	345	98.01	7	1.99
Chems	240	232	96.67	8	3.33
BusEq	928	903	97.31	25	2.69
Telcm	160	158	98.75	2	1.25
Utils	448	446	99.55	2	0.45
Shops	736	721	97.96	15	2.04
Hlth	544	531	97.61	13	2.39
Finance	1,232	1210	98.21	22	1.79
Others	758	743	98.80	9	1.20
Sum	6,432	6,306		126	

[24] Industry titles are based on four-digit Standards Industrial Classification (SIC) 12 codes provided by Fama and French.

Table 5
Descriptive Statistics – Number of CAPs per Year and Industry

Panel A: By Year								
	Mean	**Min**	**25 %**	**Median**	**75 %**	**Max**	**SD**	**N**
2001	3	0	0	3	4	11	2	1,107
2002	5	0	3	5	6	13	2	1,880
2003	5	0	4	5	6	12	2	2,036
2004	5	1	4	5	6	12	2	2,128
2005	5	1	4	5	7	12	2	2,170
2006	6	1	4	5	7	13	2	2,297
2007	6	1	4	6	7	13	2	2,357
2008	6	2	5	6	7	14	2	2,422
2009	6	1	5	6	7	14	2	2,460
2010	6	2	5	6	7	14	2	2,478
2011	6	1	5	6	7	14	2	2,446
2012	6	2	5	6	7	14	2	2,430
2013	6	1	5	6	7	14	2	2,419
2014	6	1	5	6	7	14	2	2,377
2015	6	1	4	6	7	14	2	2,344
2016	6	1	4	6	7	14	2	2,335
Σ	6	0	4	5	7	14	2	35,686
Panel B: By Industry								
NoDur	6	0	5	6	7	12	2	2,246
Durbl	6	0	5	6	7	11	2	756
Manuf	6	0	5	6	8	13	2	3,262
Enrgy	6	0	4	5	7	11	2	1,979
Chems	5	0	4	5	6	11	2	1,238
BusEq	6	0	4	6	8	13	2	5,627
Telcm	5	0	4	5	6	8	1	835
Utils	6	0	4	5	7	11	2	2,545
Shops	6	0	4	6	7	14	2	4,368
Hlth	6	0	5	6	7	13	2	3,229
Finance	5	0	3	5	6	9	2	5,623
Others	5	0	4	5	7	10	2	3,978
Σ	6	0	4	5	7	14	2	35,686

As can be seen, all firms in my final sample provide CAP disclosures since 2004. In comparison, only one firm in 2003, seven firms in 2002 and 118 firms in 2001 did not provide any discussion of their CAPs.

Immediately after the SEC issued FR-60 in 2001, 284 companies (about 70 percent) provided on average three accounting topics (in sum 1,107 CAPs) as critical. In the following years, the average (total) number of CAPs steadily increased to six (2,478) in 2010 and

remains similar in subsequent years (2,335 CAPs in 2016).[25] There are huge discrepancies in the number of CAPs across firms and industries. While some firms only have one or two CAPs, there are others with 14 highly uncertain and complex accounting topics. Moreover, most firms in the sample are from the finance sector, followed by business equipment, retail shops and manufacturing. Firms from the telecommunication (finance) sector have the lowest number of CAPs, with a mean of five (five) and a maximum of eight (nine) CAPs. Overall, by looking at other industries, the average number of CAPs is between five and six, with a minimum of zero and a maximum between eleven and 14. This variation remains equal over time and across industries.

Table 6 provides the total number of observations as well as firms that flag an accounting topic as critical. Deferred taxes, intangibles, property, plant and equipment, retirement benefits, and revenue recognition are the five most frequent CAPs, followed by contingencies, inventories, stock-based compensation, and receivables. Accounting topics such as equity, cash and cash equivalents, commitments, and foreign currencies are less common. Furthermore, about 90 percent of my sample firms flag deferred taxes and intangibles at least once as critical, followed by property, plant and equipment and revenue recognition. Other CAPs such as receivables, investments, inventories, retirement benefits, and contingencies occur in about 50 percent of the sample. To sum up, all accounting topics (except deferred revenue) are classified as highly uncertain by at least one firm between 2001 and 2016. This demonstrates the variation of highly uncertain accounting topics across firms. The most frequent highly uncertain accounting topics are in line with the evolving financial statement complexity literature (e.g., Chychyla et al. 2018; Filzen and Peterson 2015). In addition, I present the number of observations that flag an accounting topic relating to the coding system as critical for each year (industry) in Appendix B (Appendix C).

[25] It is possible that ambiguity in CAP headings actually led to more CAPs than headings. According to my pre-test, I identified several firms with two or more accounting positions in a single heading. As a result, I manually went through the data to identify headings with multiple CAPs and coded them as single CAPs.

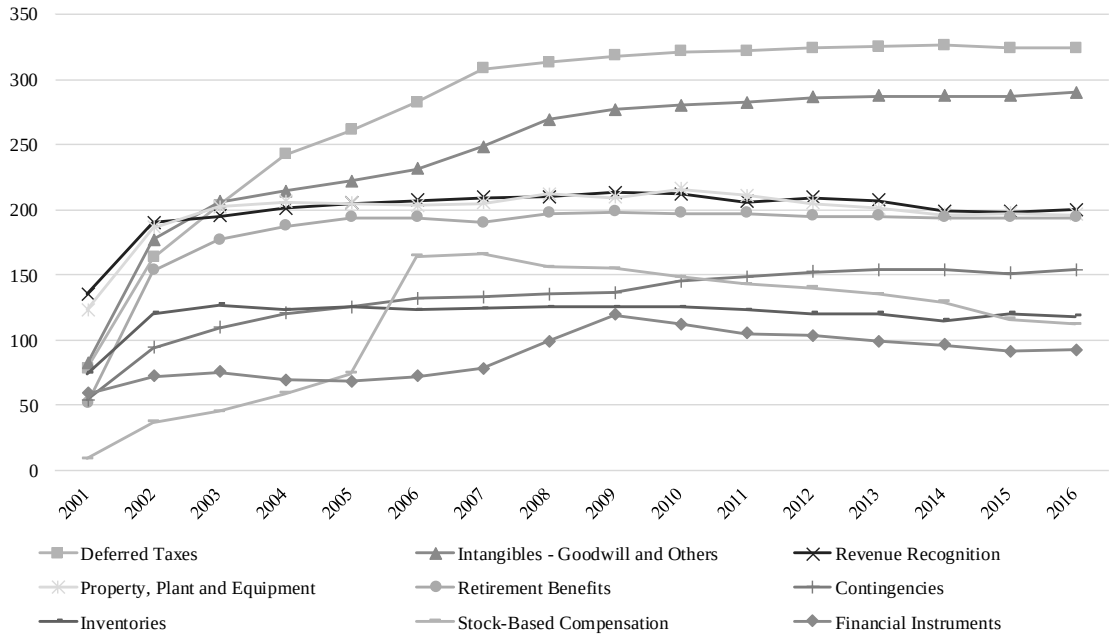
Table 6
Number of CAPs per Accounting Topic

Category	No. of Unique Firms	%	No. of Observations	%
Deferred Taxes	353	87.81	4,434	69.91
Intangibles - Goodwill and Others	344	85.57	3,926	61.90
Revenue Recognition	268	66.67	3,196	50.39
Property, Plant and Equipment	296	73.63	3,173	50.03
Retirement Benefits	224	55.72	2,907	45.84
Contingencies	215	53.48	2,096	33.05
Inventories	158	39.30	1,907	30.07
Stock-Based Compensation	205	51.00	1,788	28.19
Receivables	165	41.04	1,635	25.78
Other Assets and Deferred Costs	151	37.56	1,500	23.65
Investments	183	45.52	1,471	23.19
Financial Instruments	159	39.55	1,411	22.25
Other Expenses	98	24.38	789	12.44
Asset Retirement / Environmental	78	19.40	759	11.97
Debt	86	21.39	757	11.94
Business Combination	127	31.59	710	11.20
Guarantees	67	16.67	630	9.93
Liabilities	76	18.91	561	8.85
Regulatory Accounting	29	7.21	356	5.61
Consolidation	53	13.18	320	5.05
Leasing	48	11.94	315	4.97
Exit or Disposal Cost Obligations	53	13.18	307	4.84
Oil and Gas Accounting	21	5.22	255	4.02
Commitments	23	5.72	154	2.43
Foreign Currency Matters	25	6.22	115	1.81
Research and Development	16	3.98	114	1.80
Cash and Cash Equivalents	9	2.24	39	0.61
Interest	13	3.23	38	0.60
Equity	7	1.74	23	0.36
Deferred Revenue	0	0.00	0	0.00
Σ	402		6,432	

Below, I take a closer look at nine CAPs over time (Figure 4). The number of all CAPs presented in Figure 4 increased in the years immediately after FR-60. Subsequently, the number of firms that classified their deferred taxes, intangibles or contingencies increased further, whereas the number of CAPs that relate to retirement benefits, property, plant and equipment, as well as revenue recognition does not change at all. A look at the curve of stock-based compensation and financial instruments reveals a significant increase in 2006 and 2008/2009, respectively. The former might originate from an accounting change

regarding stock-based compensation in 2005 (Statement of Financial Accounting Standards (SFAS) 123 to SFAS 123R). This change led to material changes with respect to the recognition of share-based payments (Frederickson et al. 2006). The latter one might be attributable to the global financial crisis between 2007 and 2009. Thus, it does not seem particularly surprising that there is an increased number of firms classifying their financial instruments (including fair value accounting) as highly uncertain. However, both curves decrease slightly afterwards. To sum up, whereas most CAPs occur consistently between 2001 and 2016, there are others with slight changes over time.

Figure 4
Occurrence of Nine CAPs Between 2001 and 2016



Lastly, to evaluate differential disclosures practices, I provide disclosure frequencies per year (Panel A) and industry (Panel B) in Table 7. Whereas the most frequent CAPs only changed slightly, there is huge variation across industries. For instance, CAPs regarding investment occurs mainly in firms of the financial sector, whereas uncertain oil, gas and regulatory accounting topics appear more frequently in firms from the energy and utility industry. Moreover, retailers (shops) tend to flag inventories as critical, while firms from the business equipment and healthcare industry mostly have CAPs relating to revenue recognition. However, while other CAPs occur across all industries, I follow Levine and Smith (2011) and argue that there still might be industry- or firm-specific judgements and uncertainties in the application of related accounting policies.

Table 7

Top Three CAPs by Year and Industry

	Most Frequent CAPs		
Panel A: By Year	**1st**	**2nd**	**3rd**
2001	Revenue Recognition	Property, Plant and Equipment	Intangibles – Goodwill and Others
2002	Revenue Recognition	Property, Plant and Equipment	Intangibles – Goodwill and Others
2003	Intangibles – Goodwill and Others	Deferred Taxes	Property, Plant and Equipment
2004	Deferred Taxes	Intangibles – Goodwill and Others	Property, Plant and Equipment
2005	Deferred Taxes	Intangibles – Goodwill and Others	Property, Plant and Equipment
2006	Deferred Taxes	Intangibles – Goodwill and Others	Revenue Recognition
2007	Deferred Taxes	Intangibles – Goodwill and Others	Revenue Recognition
2008	Deferred Taxes	Intangibles – Goodwill and Others	Property, Plant and Equipment
2009	Deferred Taxes	Intangibles – Goodwill and Others	Property, Plant and Equipment
2010	Deferred Taxes	Intangibles – Goodwill and Others	Property, Plant and Equipment
2011	Deferred Taxes	Intangibles – Goodwill and Others	Property, Plant and Equipment
2012	Deferred Taxes	Intangibles – Goodwill and Others	Revenue Recognition
2013	Deferred Taxes	Intangibles – Goodwill and Others	Revenue Recognition
2014	Deferred Taxes	Intangibles – Goodwill and Others	Revenue Recognition
2015	Deferred Taxes	Intangibles – Goodwill and Others	Revenue Recognition
2016	Deferred Taxes	Intangibles – Goodwill and Others	Revenue Recognition
NoDur	Deferred Taxes	Intangibles – Goodwill and Others	Retirement Benefits
Durbl	Retirement Benefits	Deferred Taxes	Guarantees
Manuf	Deferred Taxes	Retirement Benefits	Intangibles – Goodwill and Others
Enrgy	Oil and Gas Accounting	Deferred Taxes	Retirement Benefits
Chems	Retirement Benefits	Deferred Taxes	Intangibles – Goodwill and Others
BusEq	Revenue Recognition	Deferred Taxes	Intangibles – Goodwill and Others
Telcm	Deferred Taxes	Intangibles – Goodwill and Others	Property, Plant and Equipment
Utils	Retirement Benefits	Regulatory Accounting	Property, Plant and Equipment
Shops	Inventories	Deferred Taxes	Property, Plant and Equipment
Hlth	Revenue Recognition	Deferred Taxes	Intangibles – Goodwill and Others
Finance	Intangibles – Goodwill and Others	Investments	Deferred Taxes
Others	Deferred Taxes	Intangibles – Goodwill and Others	Property, Plant and Equipment

6.2 Reporting Characteristics

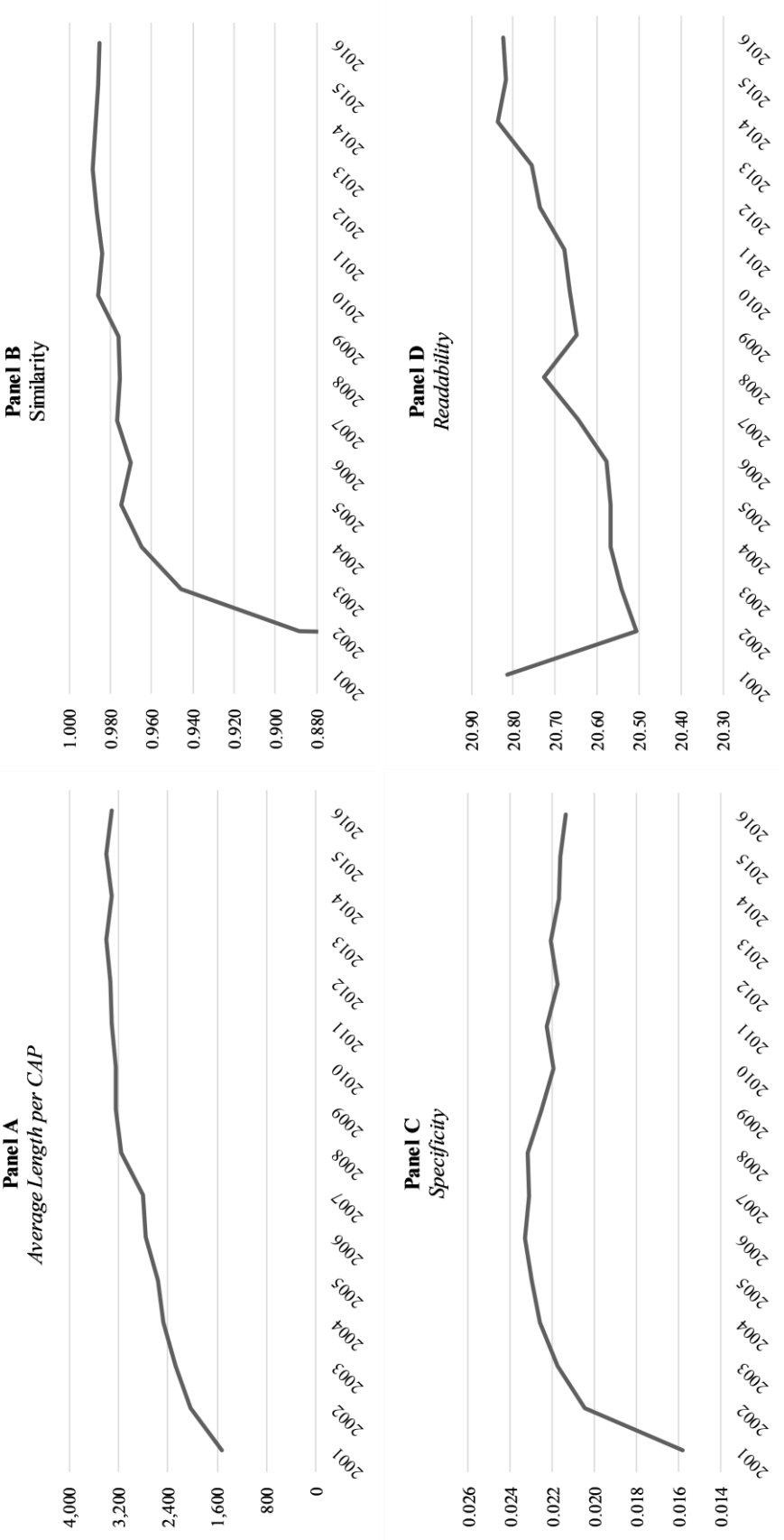
Research question 2 focuses on the characteristics of CAP disclosures. Figure 5 depicts the mean of all four textual attributes over time and Table 8 presents detailed descriptive statistics.

Table 8
Descriptive Statistics – Reporting Characteristics

Panel A: Reporting Characteristics - Pooled										
	n	Mean	Min	p5	p25	p50	p75	p95	max	sd
Length per CAP	6,308	2,907	95	894	1,668	2,411	3,521	6,459	26,162	2,073
Length per Section	6,308	15,612	192	3,880	8,732	13,703	19,573	33,694	106,569	10,594
Similarity	5,906	0.973	0.511	0.896	0.971	0.989	0.996	0.999	1.000	0.044
Specificity	6,308	0.022	0.000	0.006	0.014	0.021	0.029	0.041	0.097	0.011
Readability	6,308	20.68	14.73	18.51	19.81	20.64	21.46	23.09	30.16	1.400

Panel B: Average Textual Attributes per Year					
	Length per CAP	Length per Section	Similarity	Specificity	Readability
2001	1,516	5,334	.	0.016	20.81
2002	2,039	9,027	0.888	0.020	20.51
2003	2,266	10,664	0.945	0.022	20.54
2004	2,467	12,136	0.965	0.023	20.57
2005	2,553	12,864	0.975	0.023	20.57
2006	2,746	14,554	0.970	0.023	20.58
2007	2,798	15,247	0.977	0.023	20.65
2008	3,153	17,642	0.975	0.023	20.73
2009	3,252	18,623	0.976	0.023	20.65
2010	3,243	18,842	0.986	0.022	20.67
2011	3,300	18,866	0.984	0.022	20.68
2012	3,324	18,919	0.987	0.022	20.74
2013	3,405	19,112	0.988	0.022	20.76
2014	3,311	18,381	0.987	0.022	20.84
2015	3,395	18,418	0.986	0.022	20.82
2016	3,319	18,067	0.986	0.021	20.83
Panel C: Average Textual Attributes per Industry					
NoDur	2,592	15,103	0.975	0.026	20.26
Durbl	2,184	13,014	0.964	0.025	20.23
Manuf	2,397	14,516	0.974	0.023	20.57
Enrgy	2,958	15,962	0.979	0.020	20.62
Chems	2,575	13,305	0.980	0.027	20.93
BusEq	2,696	15,475	0.976	0.019	20.67
Telcm	2,658	13,616	0.971	0.028	20.85
Utils	2,857	16,259	0.966	0.027	20.98
Shops	1,921	11,611	0.973	0.020	20.32
Hlth	2,582	15,485	0.972	0.020	20.73
Finance	4,097	19,116	0.973	0.020	21.07
Others	3,210	16,105	0.972	0.023	20.49

Figure 5
Reporting Characteristics of CAP Disclosures



6.2.1 Length

Figure 5 (Panel A) and Table 8 depict the average length per CAP between 2001 and 2016. Most notably, the average length per CAP increased significantly from about 1,500 in 2001 to 2,000 characters in 2002.[26] This may be attributable to the missing disclosure guidelines in 2001 and the detailed Proposed Rule in 2002. Subsequently, the average length of each CAP increased substantially and by roughly the same rate to about 3,400 characters in 2013, and decreased slightly to 3,300 characters in 2016. Table 8 reveals that there is substantial variation in the number of characters disclosed by firms. Whereas the fifth (95th) percentile is 894 (6,459) characters, the median (standard deviation) is 2,411 (2,073) characters. Additionally, while some firms have complete CAP sections with more than 90,000 characters (e.g., Hartford Financial Service Group Inc. 2009), there are others with less than 2,000 characters (e.g., Warner Media LLC 2009) or 400 characters (e.g., Progressive Corporation 2009). Moreover, there is also a constant increase in the total number of characters per CAP section. The average length is about 11,000 characters in 2003 and increases constantly to about 18,000 characters in 2016, with a maximum of more than 19,000 characters in 2013. By considering each year separately, I find that the huge variation in the average length of each CAP, as well as the CAP section, remains unchanged.

6.2.2 Textual Similarity

Results with respect to the textual similarity of CAP disclosures over time are presented in Panel B of Figure 5 and Table 8. There are no values in 2001, because there were no disclosures in 2000. Whereas the average cosine similarity is 0.888 in 2002, it increases constantly to 0.986 in 2016. This indicates that CAP disclosures are quite similar over time and the year-to-year change in the content is up to only two to three percent on average. Table 8 shows the variation of the cosine similarity. Note that there is modest variation with scores of 0.896 (0.999) at the fifth (95th) percentile and 0.971 (0.996) at the 25th (75th) percentile. Thus, the application and occurrence of highly complex accounting policies and estimates for each firm separately, as well as the related disclosure does not change substantially over time. Moreover, the standard deviation is relatively

[26] One page contains on average 3,000 characters.

low (0.044), which can be expected, since the occurrence of highly complex accounting topics should be relatively stable over time.

6.2.3 Specificity

The mean (median) value of specificity is 0.022 (0.021), suggesting that on average, two out of 100 words are specific (Table 8). As in the case of similarity, I find slight variation for specificity over time (Panel C of Figure 5). On the one hand, the 25th (75th) percentile implies that about one (three) out of 100 words are specific. On the other hand, whereas some CAP sections do not contain any specific information, there are others with a score of 0.097, indicating that 10 out of 100 words are specific. Most importantly, whereas the specificity of CAP disclosures increases slightly before 2008, there is a continuous decrease in subsequent years. Comparing my results with prior findings, Hope et al. (2016) find that five out of 100 words are specific by analyzing risk-factor disclosures.[27]

6.2.4 Readability

As can be seen in Panel D of Figure 5 and Table 8, most of the CAP sections are on average very difficult to understand (i.e., Fog index above 18). The mean (median) of the Fog index for the entire CAP section is about 20.7 (20.6) that indicates 'unreadable' according to the interpretation of the index (e.g., Li 2008). Comparing my findings with prior literature, Lang and Stice-Lawrence (2015) find that the average readability score of annual reports out of 42 countries is about 19.5 (for similar results see Li (2008)). Therefore, CAP sections are on average slightly less readable than the remaining 10-K. In addition, Panel D of Figure 5 shows a sharp decrease in the readability score in 2002, indicating CAP sections that are more readable. Nonetheless, the score increases again constantly in subsequent years, implying less readable disclosures. However, these changes are negligible, because overall, CAPs remain 'unreadable'. The standard deviation is 1.4, indicating that there is only modest variation. According to the interpretation of the index, only a few observations (< 5 %) have CAP sections that are 'difficult' (readability score < 18) and no CAP section is 'easy' to read (readability score < 14). Assuming that the complexity of the underlying accounting topics involve the use of complex

[27] Using the Standford NER tool instead of the SpaCy toolkit, I identify between one and two specific words out of each 100 words. However, the difference between both toolkits may be attributed to the smaller number of categories.

and technical language, it is not surprising that CAP disclosures are on average unreadable.

6.3 Compliance with the SEC Guidance

Research question 3 asks whether CAP disclosures comply with the SEC guidelines. Herewith, I analyze CAP disclosures of the largest 100 firms in 2016 using content analysis.

Table 9 breaks down the content of each CAP by the main topics of the SEC guidelines.[28] In sum, the sample of 100 firms disclosed 569 CAPs in 2016. The most frequent topics are deferred taxes (80 firms), retirement benefits (72), intangibles, goodwill and others (60), and property, plant and equipment (60). The results in Table 9 show that 95 firms provide a general description at the beginning of each CAP section and 24 firms already present a short overview of their CAPs. Moreover, 52 firms include a reference to related notes or financial line items for further information.

By looking separately at each subject in the SEC guidance, firms provide, for 98 percent of their CAPs (556 out of 569), a description of the methodology and for 56 percent (320), general information regarding the nature of the estimate. Moreover, about 50 percent of all CAP disclosures (261) include a discussion of the underlying assumptions needed for the estimate, the overall result of the estimate (276), and its impact on financials (293). Nearly 70 percent of all CAPs (392) contain information about the factors that affect the method and/or assumption. Furthermore, firms include a sensitivity analysis providing quantitative information with respect to the impact of a change in the estimate for 156 of their CAPs (27 %). Most of them relate to retirement benefits (65), property, plant and equipment (16), intangibles (14), asset retirement (9), and revenue recognition (8). Compared to that, very few firms provide quantitative information on the impact of the estimate on financials (eight firms for nine CAPs) and the accuracy of the estimate in the past (ten firms for twelve CAPs). Whereas 37 firms explain the reasons for undertaken changes to the estimates of 47 CAPs (8.3 %) in the past years, only 30 CAPs include a quantitative discussion of these changes.

[28] See Table 1 for an overview of the required subjects mentioned in the SEC releases. I further aggregate two additional subjects: Result of the estimate as well as references to the notes of the financial statements.

Table 9

Disclosure Compliance with SEC Guidance - Number of Firms Making Disclosures about a Given Subject and Accounting Topic

	Accounting Topic																																				
Subject of CAP Disclosure	Introductory Information	Cash and Cash Equivalents	Receivables	Investments	Inventories	Other Assets and Deferred Costs	Intangibles – Goodwill and Others	Property, Plant and Equipment	Liabilities	Asset Retirement / Environmental	Exist or Disposal Cost Obligations	Deferred Revenue	Commitments	Contingencies	Guarantees	Debt	Equity	Revenue Recognition	Retirement Benefits	Stock Compensation	Other Expenses	Research and Development	Deferred Taxes	Business Combinations	Consolidation	Financial Instruments	Foreign Currency Matters	Interest	Leasing	Regulatory Accounting	Oil and Gas Accounting	No. of Firms	Σ	Prop.
Introductory Information																																					
General Description	95																															95		
CAP Listing	24																															24		
Nature of the Estimate			3	6	5	14	24	24	4	12	2		1	41	1	2		30	48	1	6		57	5	5	10	1	3	11	2	2	97	320	0.562
Methodology			11	15	18	18	58	59	5	25	3		1	48	6	5		44	70	6	13	1	78	14	6	19	2	7	16	4	4	98	556	0.977
Assumptions	1		2	4	4	8	36	36	1	14				8	1	2		13	67	3	3	1	28	12	1	9		2	2	2	2	94	261	0.459
Factors Affecting the Methodology	3		8	14	10	15	43	49	3	21	2		1	27	4	3		29	51		6		57	8	3	12	2	6	12	3	3	98	392	0.689
Result of the Estimate	1		5	3	7	10	49	35		13				15	2	3		22	39	1	4		41	1	2	5	1	3	7	4	4	94	276	0.485
Impact of Estimate on Financials																																					
Narrative Information	5		5	8	9	6	33	30	2	12			1	29	3	2		21	36	2	4		47	11	2	9	1	5	13	1	1	96	293	0.515
Quantitative Information							1			1									3		2		2									8	9	0.016
Accuracy of the Estimate																																					
Narrative Information	1				1	2		2		1								4					1			1						10	12	0.021
Quantitative Information																																					
Changes to Estimate in Past Years																																					
Narrative Information					1	1	2	2		1				1				3	28	2	1		2	1				1	1			37	47	0.083
Quantitative Information						1		3		2								1	19	1			2						1			27	30	0.053
Sensitivity Analysis	1		3	4	3	5	14	16	2	9				1	2	4		8	65	1			5	1		1		6		3	3	81	156	0.274
Discussion with the Audit Committee	34																															34		
Discussion on a Segment Basis							5	2										2	1	1												10	11	0.019
Reference to Notes	52		1	5	2	4	24	21	3	18	2		1	36	1	2		11	54	3	4		37	9	2	16		3	13	3	3	95	278	0.489
No. of Firms	95	0	11	17	19	18	60	60	5	26	3	0	1	52	6	5	0	45	72	6	13	1	80	14	6	20	2	0	7	16	4		569	

In addition, eleven CAPs (1.9 %) were discussed on a segment basis, and only 34 firms discussed the selection of their CAPs with the audit committee.

Table 10 presents the average proportion of each subject (estimated by the number of characters) compared to the length of each CAP.[29] As can be seen, information regarding the underlying methodology captures about 50 percent of the presented disclosures, whereas on average, 20 percent of the content relates to the nature of the estimate, the underlying assumptions, factors affecting the method and/or assumption, as well as quantitative and qualitative information with respect to changes made to the estimate. Subsequently, 12.5 percent of the content contain information regarding the impact of the estimate on other financial statements and 8.6 percent relate to the past accuracy of the estimate. A quantitative discussion of the sensitivity of each CAP comprises about 12.4 percent, with a minimum of 1.6 percent and a maximum of 52.4 percent and is highest for topics that relate to retirement benefits, stock-based compensation, inventories, contingencies, and liabilities.

Table 10
Descriptive Statistics – Proportion of each CAP Subject

	Mean	**Min**	**25 %**	**Median**	**75 %**	**Max**	**SD**
Nature of the Estimate	0.175	0.012	0.074	0.141	0.234	0.744	0.136
Methodology	0.478	0.033	0.335	0.466	0.610	1.000	0.211
Assumptions	0.162	0.010	0.088	0.133	0.207	0.618	0.110
Factors affecting the Methodology	0.196	0.017	0.097	0.172	0.263	0.650	0.123
Result of the Estimate	0.152	0.007	0.080	0.132	0.192	0.551	0.101
Impact on Financials							
Narrative Information	0.125	0.010	0.053	0.100	0.168	0.748	0.100
Quantitative Information	0.079	0.027	0.041	0.067	0.118	0.157	0.048
Accuracy of the Estimate							
Narrative Information	0.086	0.007	0.044	0.062	0.082	0.390	0.096
Quantitative Information	-	-	-	-	-	-	-

[29] Appendix D depicts the average length of each accounting topic and Appendix E presents the average proportion of each subject compared to the length of each accounting topic.

Table 10 - continued
Descriptive Statistics – Proportion of each CAP Subject

	Mean	Min	25 %	Median	75 %	Max	SD
Changes in Past Years							
Narrative Information	0.135	0.014	0.059	0.107	0.207	0.372	0.097
Quantitative Information	0.082	0.015	0.035	0.066	0.109	0.223	0.060
Sensitivity Analysis	0.124	0.016	0.061	0.105	0.158	0.524	0.093
Discussion with the Audit Committee	0.167	0.067	0.111	0.155	0.202	0.324	0.069
Discussion on a Segment Basis	0.056	0.002	0.017	0.039	0.080	0.235	0.065
Reference to the Notes	0.102	0.003	0.042	0.071	0.117	1.000	0.116

Having said that, the main part of each CAP discussion captures information on the methodology used to determine the estimate, followed by general information about the estimate, the underlying assumptions, and factors affecting the methodology and assumptions. By contrast, firms provide less information with respect to the impact of the estimate on financials, the accuracy of the estimate in the past, and changes made to the estimate in past years. However, this finding is not surprising, since the SEC only requires such disclosures if applicable. Other subjects such as a statement on whether the firm has discussed the selection of their CAPs with the audit committee, or references to the notes, are mostly presented in a short statement. In addition, while some firms discuss almost all subjects required by the SEC, there are others that only describe the methodology used to determine the estimate (e.g., AbbVie Inc. 2016; Alphabet Inc. 2016; Merck & Co. Inc. 2016) or refer only to the related notes, without providing any further information (e.g., Mondelez International 2016; Pfizer Inc. 2016).

7 Discussion, Practical Implications and Future Research Suggestions

Over the years, the SEC has continued to focus on the regulatory framework, as well as improvements to CAP disclosures, especially in the first few years after 2001. Despite all regulatory efforts to mandate CAP disclosures, no final rule has been published to this day. Nevertheless, the percentage of companies providing CAP disclosures increased from 70 percent in 2001 to 100 percent in 2004. This finding suggests that the SEC's recommendations to disclose CAPs have a quasi-statutory binding effect for companies. In their initial releases, the SEC states that the number of CAPs will be about three to five

CAPs and should vary between companies (Levine and Smith 2011; SEC 2002a). I find that firms flag on average six accounting topics as highly uncertain. Thus, firms provide slightly more CAPs than the SEC suggested. Whereas my results show a significant increase in the number of CAPs between 2001 and 2007, there are only slight changes in the years thereafter. This finding might be attributed to the evolving and in part, unspecific guidelines provided by the SEC, as well as to the fact that firms initially had to learn how to implement the new regulation. Furthermore, the identified accounting topics that are flagged as CAPs vary between firms and industries. This finding is in line with the SEC's intention that each firm should flag those accounting policies and estimate that are most important and representative for their business (e.g., SEC 2002a). However, while some CAPs are more common (i.e., retirement benefits, property, plant and equipment, and intangibles), there might be still industry- or firm-specific judgements and uncertainties in the application of related accounting policies.

In contrast to the assumption that the occurrence of highly uncertain accounting policies might usually be stable over time, I find that the number of uncertain accounting topics and the total number of CAPs vary (slightly) from year to year. This finding raises the question as to what determines a firm's decision to flag an accounting topic as critical. First, I provide some exploratory evidence that accounting-standard changes, as well as macroeconomic developments, might affect the number of single complex accounting topics. This assumption is consistent with the idea that firms should communicate the effects of specific trends, events, and uncertainties on their methods, assumptions, and estimates used to determine their financial statements (SEC 2001). Second, firms might be confronted with an overall increasing uncertainty at the market, business, and accounting levels that determine the uncertainty associated with new and/or existing accounting topics. However, analyzing the main drivers is difficult, because the overall uncertainty in financial statements depends upon a complex entanglement of several factors at the organizational, market, business, and accounting levels (e.g., Filzen and Peterson 2015; Kuschel 2015). As shown in the review of prior studies, only some studies analyze the determinants of CAP disclosures. However, none focuses on macroeconomic developments or accounting-standard changes. While the aforementioned findings are exploratory in their nature, more research is needed to expand and deepen our understanding of the factors that determine the decision to flag accounting policies and their estimates as highly uncertain.

With respect to my second research question, I use tools from computational linguistics to characterize the textual data provided in a firm's CAP section across a number of distinct dimensions. I find that the average length of the complete section, as well as for each CAP, has constantly increased over time and that there is great variation between firms (Panel A of Figure 5 and Table 8). In this context, the SEC emphasizes that companies should not provide a "lengthy discussion of a multitude of accounting estimates in which the truly critical ones are obscured" (SEC 2002a, p. 12). Nevertheless, it is not clear whether firms with longer CAP sections also provide disclosures that are more informative. Whereas prior studies argue that longer disclosures tend to be more informative (e.g., Lang and Stice-Lawrence 2015), it is possible that firms may still try to obscure relevant information either by providing non-disclosures or lengthy discussions of their CAPs (SEC 2002a). Future research might focus on this issue and analyze whether related disclosures are to a large extent boilerplate information, or firm-specific with respect to measurement uncertainties in accounting estimates. Furthermore, CAP disclosures provided by each firm are very similar over time. However, there are ambiguous interpretations of similar CAP disclosures. On the one hand, providing similar disclosures from year-to-year might imply consistency in the occurrence of uncertainties in the measurement process. In this context, the FASB argues that consistency is an important aspect of financial reporting (FASB 2010; Peterson et al. 2015). Consequently, financial statement users may be used to the information provided by CAPs that might resolve, to some extent, the degree of uncertainty about the future (Bozanic and Thevenot 2015). On the other hand, similar disclosures might also be interpreted as redundant information that had been disclosed before, thus indicating no additional information content. Accordingly, it remains an empirical question analyzing the effects of similar CAP disclosures. Besides, CAP disclosures are on average complex and do not generally contain specific information. On the one hand, CAP disclosures that are more specific might enable investors to interpret more accurately the uncertainties within the measurement process of financial statements, whereas less readable information might be more difficult to understand. It seems that my descriptive findings might contrast with the SEC intention to provide understandable, clear, and specific disclosures in plain English. Nonetheless, to this day, it remains unknown whether the aforementioned characteristics are really associated with (negative) consequences.

While quantitative information about CAP disclosures and their association with firm and economic outcomes has been studied before, prior research has failed to analyze the consequences of the qualitative information beyond CAP disclosures. Yet, this is important, because one might expect them to be correlated with the information content, and thus to how outsiders perceive related disclosures (Lang and Stice-Lawrence 2015). Future research could assess how the aforementioned attributes are associated with economic outcomes (i.e., information asymmetry) or analyst data (i.e., analyst coverage, analyst forecasts, and analyst dispersion) in order to contribute to the current debate about the usefulness of CAP disclosures. Furthermore, it remains unknown whether CAPs actually capture uncertainties embedded in accounting positions. In this context, CAPs might simply reflect hypothetical uncertainties, so that firms are only minimally compliant with the SEC guidelines, so as to avoid further scrutiny (Glendening 2017) or litigation (Levine and Smith 2011). In this context, future research might focus on whether CAPs actually reflect instances of greater measurement uncertainty embedded in financial statement positions. For instance, in accordance with prior literature (e.g., Barth et al. 2016; Barth et al. 2001; Dechow and Dichev 2002), one might analyze whether accruals that are flagged as CAPs still have predictive value with respect to future cash flows or earnings.

FR-60, the Proposed Rule, as well as FR-72, still serve as major guidelines in the preparation of CAP disclosures. However, the content of each CAP section is still left to the discretion of each firm, due to the absence of a legal requirement. In their 2003 review of Fortune 500 firms, the SEC identified 14 accounting topics that could be more transparent with respect to the discussion of estimates and assumptions. Hughes et al. (2009) find an increased trend towards compliance with the SEC guidance from 2001 to 2003. Although he finds some slight improvements, various subjects mentioned in FR-72 remain underdisclosed and most firms still did not respond to the specific guidance included in the Proposed Rule. Compared to their study, my findings show that CAP disclosures in 2016 have improved qualitatively, as estimated by the level of compliance with the SEC guidance, as well as the breadth and depth of each CAP. I find that almost all firms that disclose changes to their estimates also provide a discussion including quantitative elements of these changes. Moreover, the majority of CAPs contain information with respect to the methodology, assumptions, as well as factors affecting the underlying assumptions and methodology. This allows financial statement users to obtain in-depth knowledge about the measurement and underlying assumptions of each CAP, the various factors included

in the estimates, as well as being able to predict future changes in the estimate and the related effects on a firm's financial condition.

My results further reveal that there are still some discrepancies in compliance with the SEC guidelines across the accounting topics. On the one hand, firms refer to almost all subjects required by the SEC when discussing CAPs that relate to inventories, other assets and deferred costs, intangibles, property, plant and equipment, asset retirement, revenue recognition, retirement benefits, and deferred taxes. On the other hand, disclosures about stock compensation, research and development, foreign currency matters, commitments, and liabilities comply at least with the requirements. Having said that, it seems that firms provide disclosures that are more comprehensive for those CAPs which occur more frequently. About 98 percent of my sample firms disclose information with respect to the nature of the estimate, the methodology and assumptions used to determine the estimate, factors affecting the method and/or assumption, as well as the result of the estimate. Furthermore, 81 percent of the firms provide a quantitative sensitivity analysis for at least one of their CAPs. Also, most firms (96 %) disclose a discussion of the impact of the estimate on their financials for one of their CAPs. The other subjects only appear occasionally across all CAPs.

Despite the trend towards increased compliance with the SEC guidance, some subjects remain underdisclosed. Although the number of firms that discuss their CAPs with the audit committee has increased, the majority of firms do not. This is contrary to the SEC's efforts, because the SEC still suggests discussing the identification and discussion of their CAPs with the audit committee, so as to improve the quality and transparency of related disclosures (Hughes et al. 2009; SEC 2002a). In this context, the SEC assumes that a discussion with the audit committee would give investors greater insight into the reliability of a firm's reported earnings and overall financial performance (SEC 2002a). As a result, firms that still do not discuss their CAPs with the audit committee should do so. Moreover, it is also problematic that the majority of firms provide information with respect to the impact of the estimate on financials, whereas only a few firms include a quantitative discussion. Providing investors with such information may improve their understanding of the extent to which other financial statements may be affected by CAPs, as well as the interplay of uncertainties with the measurement of financial statements. Nonetheless, there are still huge discrepancies across the subjects mentioned in the releases of the SEC. Despite increased compliance with the SEC guidelines, there seems to be still

room for improvement. Thus, the SEC should be attentive to any missing subjects when reviewing a firm's CAP disclosures, for instance, in their filing review process (e.g., Cassell et al. 2013).

8 Conclusion

In the early 2000s, the SEC issued guidance on the disclosure of CAPs, so that financial statement users should be able to assess information about highly complex and uncertain accounting policies and estimates that are helpful for understanding the measurement basis of financial statements. The primary research questions in this study are: (1) How have the number and most frequent topics of CAPs across firms and industries evolved over time, (2) what are the reporting characteristics of CAP disclosures and (3) how did they comply with the SEC guidelines in 2016. Whereas most prior studies focus on single years or on those immediately after FR-60, this is the first comprehensive empirical study analyzing CAP disclosures over a 16-year period from 2001 to 2016. My findings therefore offer new insights and key contributions compared to prior literature.

I conclude that CAP disclosures provide detailed insights into the distribution, variation, and occurrence of highly uncertain accounting policies and estimates across firms, industries, and over time. Compared to prior literature, I find that there is an increased compliance with SEC guidelines in 2016. Related disclosures include detailed information about measurement basis, as well as the effect of uncertainty on other financial variables that might enhance investor understanding about the reliability of accounting estimates. However, my findings are only descriptive in nature, and thus have to be interpreted with caution. Future research might use multivariate approaches to enhance our understanding of the effectiveness, usefulness, and quality of CAP disclosures. Nonetheless, my findings enhance current knowledge about the occurrence and regulatory framework of CAPs, how CAPs are reported, as well as the content provided in each CAP section. Having said that, the SEC still strives to mandate CAP disclosures. Despite all regulatory efforts, no final rule has been published and still it is questionable whether the SEC will revise Item 303 of Regulation S-K in the near future. Given an apparent increased compliance with the SEC guidance, there should be clearer principal-based guidelines for removing current regulatory gaps, so as to improve the quality and transparency of CAP disclosures.

To sum up, CAPs contribute to the SEC's goal of communicating highly uncertain accounting policies and estimates and how uncertainties inherent in accounting estimates affecting a company's financial performance.

Appendix A: Category System – CAPs

	Topic Category 1	Topic Category 2
1	Cash and Cash Equivalents	Cash and Cash Equivalents
2	Receivables	Receivables
3	Investments	Debt and Equity Securities
		Equity Method and Joint Ventures
4	Inventories	Inventories
5	Other Assets and Deferred Costs	Insurance Contracts
		Contracts With Customers
		Other Assets
6	Intangibles - Goodwill and Others	Goodwill
		Intangibles – Other than Goodwill
		Internal-Use Software
7	Property, Plant and Equipment	Property, Plant and Equipment
8	Liabilities	Liabilities
9	Asset Retirement / Environmental	Asset Retirement Obligations
		Environmental Obligations
10	Exit or Disposal Cost Obligations	Restructuring
		Exit and Closing Obligations
11	Deferred Revenue	Deferred Revenue
12	Commitments	Commitments
13	Contingencies	Contingencies
14	Guarantees	Guarantees
15	Debt	Reserves
		Loans
		Other Debt
16	Equity	Equity
17	Revenue Recognition	Revenue Recognition
		Returns
		Rebates
		Other Income
18	Retirement Benefits	Retirement Benefits
19	Stock-based Compensation	Stock Compensation
		Other Incentives
20	Other Expenses	Other Expenses
21	Research and Development	Research and Development
22	Deferred Taxes	Deferred Taxes
23	Business Combination	Business Combination
24	Consolidation	Consolidation
25	Financial Instruments	Derivatives
		Hedging
		Financial Instruments
		Fair Value Accounting
26	Foreign Currency Matters	Foreign Currency Matters
27	Interest	Interest
28	Leasing	Leasing
29	Regulatory Accounting	Regulatory Accounting
30	Oil and Gas Accounting	Oil and Gas Accounting

Appendix B: Number of Accounting Topics per Year

Name	**2001**	**2002**	**2003**	**2004**	**2005**	**2006**	**2007**	**2008**	**2009**	**2010**	**2011**	**2012**	**2013**	**2014**	**2015**	**2016**	**Total**
Cash and Cash Equivalents	1	4	2	1	1	1	2	3	3	3	3	4	4	2	3	2	**39**
Receivables	72	115	119	122	119	114	108	111	110	105	102	98	95	87	83	75	**1,635**
Investments	61	93	89	94	93	87	87	95	98	101	101	98	97	93	93	91	**1,471**
Inventories	74	120	127	123	125	123	124	125	125	125	123	120	120	115	120	118	**1,907**
Other Assets and Deferred Costs	56	100	100	103	105	105	101	102	100	98	93	92	90	89	83	83	**1,500**
Intangibles – Goodwill and Others	83	177	206	214	222	231	248	269	277	280	282	286	287	287	287	290	**3,926**
Property, Plant and Equipment	123	186	202	206	205	203	205	212	209	216	211	205	201	196	197	196	**3,173**
Liabilities	25	36	32	36	33	36	36	36	32	35	35	37	39	39	38	36	**561**
Asset Retirement	17	35	44	42	44	45	52	53	56	55	52	53	53	53	53	52	**759**
Exit or Disposal Cost Obligations	14	31	27	24	24	19	21	18	20	18	15	14	16	15	15	16	**307**
Deferred Revenue	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	**0**
Commitments	10	12	9	8	10	9	9	10	11	11	11	10	9	8	8	9	**154**
Contingencies	54	94	109	120	125	132	133	135	136	145	148	152	154	154	151	154	**2,096**
Guarantees	17	34	40	40	37	40	41	43	44	44	42	44	44	43	40	37	**630**
Debt	38	54	54	50	46	48	48	47	49	48	47	47	46	47	44	44	**757**
Equity	1	4	2	2	0	0	0	0	1	1	2	2	2	2	2	2	**23**
Revenue Recognition	135	190	195	201	205	207	209	210	213	212	206	209	207	199	198	200	**3,196**
Retirement Benefits	51	153	177	187	194	194	190	197	198	197	197	195	195	194	194	194	**2,907**
Stock-Based Compensation	9	37	45	59	74	164	166	156	155	148	143	140	135	129	116	112	**1,788**
Other Expenses	29	56	54	56	55	54	53	51	48	51	53	48	47	46	44	44	**789**
Research and Development	4	7	9	8	8	8	9	9	7	6	7	7	8	6	6	5	**114**
Deferred Taxes	78	163	203	242	261	282	308	313	318	321	322	324	325	326	324	324	**4,434**

Appendix B: Number of Accounting Topics per Year

Name	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016	Total
Business Combination	14	19	24	25	27	34	39	42	46	54	59	57	60	67	69	74	**710**
Consolidation	11	14	18	17	16	17	21	23	21	25	23	25	24	23	21	21	**320**
Financial Instruments	61	72	75	69	68	72	78	99	119	112	105	103	99	96	91	92	**1,411**
Foreign Currency Matters	11	9	12	13	9	9	8	7	5	7	5	4	3	3	5	5	**115**
Interest	9	5	4	4	2	2	1	2	3	4	2	0	0	0	0	0	**38**
Leasing	14	23	20	23	24	23	20	17	17	18	20	18	20	20	20	18	**315**
Regulatory Accounting	22	22	23	23	22	22	22	20	22	22	22	22	22	22	23	25	**356**
Oil and Gas Accounting	13	15	15	16	16	16	18	17	17	16	15	16	17	16	16	16	**255**

Appendix C: Number of Accounting Topics per Industry

Name	NoDur	Durbl	Manuf	Enrgy	Chems	BusEq	Telcm	Utils	Shops	Hlth	Finance	Others	Total
Cash and Cash Equivalents	1	4	2	1	1	1	2	3	3	3	3	4	**39**
Receivables	72	115	119	122	119	114	108	111	110	105	102	98	**1,635**
Investments	61	93	89	94	93	87	87	95	98	101	101	98	**1,471**
Inventories	74	120	127	123	125	123	124	125	125	125	123	120	**1,907**
Other Assets and Deferred Costs	56	100	100	103	105	105	101	102	100	98	93	92	**1,500**
Intangibles – Goodwill and Others	83	177	206	214	222	231	248	269	277	280	282	286	**3,926**
Property, Plant and Equipment	123	186	202	206	205	203	205	212	209	216	211	205	**3,173**
Liabilities	25	36	32	36	33	36	36	36	32	35	35	37	**561**
Asset Retirement	17	35	44	42	44	45	52	53	56	55	52	53	**759**
Exit or Disposal Cost Obligations	14	31	27	24	24	19	21	18	20	18	15	14	**307**
Deferred Revenue	0	0	0	0	0	0	0	0	0	0	0	0	**0**
Commitments	10	12	9	8	10	9	9	10	11	11	11	10	**154**
Contingencies	54	94	109	120	125	132	133	135	136	145	148	152	**2,096**
Guarantees	17	34	40	40	37	40	41	43	44	44	42	44	**630**
Debt	38	54	54	50	46	48	48	47	49	48	47	47	**757**
Equity	1	4	2	2	0	0	0	0	1	1	2	2	**23**
Revenue Recognition	135	190	195	201	205	207	209	210	213	212	206	209	**3,196**
Retirement Benefits	51	153	177	187	194	194	190	197	198	197	197	195	**2,907**
Stock-Based Compensation	9	37	45	59	74	164	166	156	155	148	143	140	**1,788**
Other Expenses	29	56	54	56	55	54	53	51	48	51	53	48	**789**
Research and Development	4	7	9	8	8	8	9	9	7	6	7	7	**114**
Deferred Taxes	78	163	203	242	261	282	308	313	318	321	322	324	**4,434**

Appendix C: Number of Accounting Topics per Industry

Name	NoDur	Durbl	Manuf	Enrgy	Chems	BusEq	Telcm	Utils	Shops	Hlth	Finance	Others	Total
Business Combination	33	0	41	84	11	130	40	6	53	112	128	72	**710**
Consolidation	46	0	24	17	0	18	2	16	8	18	151	20	**320**
Financial Instruments	47	6	89	168	17	88	3	267	27	76	541	82	**1,411**
Foreign Currency Matters	6	0	0	25	10	2	0	11	23	4	19	15	**115**
Interest	0	0	0	4	0	0	0	0	6	0	22	6	**38**
Leasing	17	34	18	1	0	36	0	37	87	8	71	6	**315**
Regulatory Accounting	0	0	0	6	0	0	1	332	0	0	17	0	**356**
Oil and Gas Accounting	0	0	0	224	0	0	0	28	0	0	0	3	**255**

Appendix D: Average Length per Accounting Topic

Category	Average Length (# Characters)
Oil and Gas Accounting	6,395
Retirement Benefits	4,494
Property, Plant and Equipment	4,175
Intangibles - Goodwill and Others	3,829
Revenue Recognition	3,757
Asset Retirement / Environmental	3,339
Business Combination	3,165
Financial Instruments	3,049
Other Assets and Deferred Costs	2,645
Foreign Currency Matters	2,577
Deferred Taxes	2,459
Regulatory Accounting	2,434
Consolidation	2,343
Leasing	2,328
Commitments	2,004
Other Expenses	1,966
Investments	1,930
Stock-Based Compensation	1,849
Debt	1,826
Contingencies	1,740
Receivables	1,550
Guarantees	1,414
Exit or Disposal Cost Obligations	1,362
Inventories	1,354
Liabilities	1,247
Research and Development	776
Cash and Cash Equivalents	-
Deferred Revenue	-
Equity	-
Interest	-
Average Length (# Characters)	2,481

Appendix E: Proportion of each CAP Subject per Accounting Topic

	Accounting Topic																														
Subject of CAP Disclosure	Introductory Information	Cash and Cash Equivalents	Receivables	Investments	Inventories	Other Assets and Deferred Costs	Intangibles – Goodwill and Others	Property, Plant and Equipment	Liabilities	Asset Retirement / Environmental	Exist or Disposal Cost Obligations	Deferred Revenue	Commitments	Contingencies	Guarantees	Debt	Equity	Revenue Recognition	Retirement Benefits	Stock Compensation	Other Expenses	Research and Development	Deferred Taxes	Business Combinations	Consolidation	Financial Instruments	Foreign Currency Matters	Interest	Leasing	Regulatory Accounting	Oil and Gas Accounting
Introductory Information																															
General Description	0.75																														
CAP Listing	0.31																														
Nature of the Estimate			0.15	0.24	0.29	0.20	0.09	0.14	0.16	0.20	0.28		0.10	0.21	0.23	0.06		0.24	0.11	0.07	0.20	0.00	0.19	0.16	0.33	0.14	0.08		0.16	0.22	0.02
Methodology			0.48	0.49	0.58	0.43	0.51	0.54	0.45	0.34	0.47		0.54	0.46	0.56	0.44		0.55	0.31	0.59	0.56	0.88	0.50	0.49	0.57	0.57	0.41		0.41	0.45	0.62
Assumptions	0.16		0.21	0.23	0.20	0.15	0.17	0.13	0.25	0.25				0.12	0.14	0.20		0.09	0.18	0.12	0.31	0.12	0.13	0.20	0.16	0.13			0.18	0.23	0.05
Factors Affecting the Methodology	0.18		0.29	0.30	0.24	0.19	0.17	0.17	0.21	0.29	0.26		0.15	0.22	0.28	0.23		0.20	0.14		0.25		0.20	0.14	0.12	0.19	0.33		0.17	0.15	0.18
Result of the Estimate	0.14		0.21	0.29	0.17	0.11	0.17	0.15		0.10				0.14	0.17	0.17		0.13	0.15	0.38	0.20		0.15	0.07	0.11	0.16	0.33		0.21	0.15	0.10
Impact of Estimate on Financials																															
Narrative Information	0.17		0.20	0.14	0.20	0.14	0.11	0.10	0.18	0.10			0.11	0.21	0.16	0.06		0.07	0.07	0.23	0.11		0.14	0.14	0.06	0.12	0.10		0.08	0.14	0.03
Quantitative Information							0.04			0.05									0.05		0.14		0.09								
Accuracy of the Estimate																															
Narrative Information	0.13				0.08	0.22		0.02		0.04								0.06					0.06			0.07					
Quantitative Information																															
Changes to Estimate in Past Years																															
Narrative Information					0.11	0.06	0.11	0.15		0.05				0.21				0.06	0.15	0.30	0.12		0.14	0.03					0.06	0.08	
Quantitative Information						0.12		0.03		0.06								0.11	0.08	0.19			0.11							0.07	
Sensitivity Analysis	0.39		0.12	0.15	0.15	0.10	0.10	0.10	0.25	0.13				0.19	0.14	0.17		0.10	0.12	0.17			0.11	0.06		0.10			0.18		0.05
Discussion with the Audit Committee	0.17																														
Discussion on a Segment Basis							0.08	0.03										0.01	0.05	0.08											
Reference to Notes	0.18		0.08	0.09	0.11	0.10	0.05	0.08	0.11	0.08	0.26		0.10	0.15	0.06	0.12		0.06	0.07	0.06	0.11		0.05	0.11	0.02	0.14			0.04	0.09	0.09

CHAPTER 2:
Critical Accounting Policy Disclosures and the Identification of Measurement Uncertainties

Abstract

The communication of measurement uncertainties in financial reporting is of increasing concern for investors, analysts, regulators, and auditors. Since 2001, U.S. firms have been encouraged to disclose their highly uncertain accounting policies and estimates, which have a material impact on how the financial condition of a firm is presented ('*critical accounting policies*', CAPs). We find that accruals which are flagged as CAPs are less persistent with respect to future cash flows. This is consistent with the SEC's intention that CAPs should capture instances of greater measurement uncertainties embedded in the underlying accrual-based measure. We argue further that the lower persistence of single uncertain accruals is not equal across firms and industries. Specifically, we find that uncertain accrual components are not less useful in predicting future cash flows *per se*; it also depends on their importance and, to a certain extent, the specificity for a given firm. Thus, we empirically document that CAP disclosures provide a suitable channel for communicating measurement uncertainties embedded in accounting estimates and financial statement positions, thus contributing to the fundamental goal of financial reporting.

This part of the thesis is a joint project with Marcus Bravidor. The paper has been accepted for presentation at the 2019 EAA Annual Congress in Paphos and the 2019 EAA Doctoral Colloquium in Larnaca. A paper version of this part is available as Rupertus and Bravidor (2019).

Acknowledgements: We gratefully acknowledge the comments and suggestions from Rolf Uwe Fülbier, Thomas Loy, Klara Lösse, Christian Mehnert, Christina Scharf, Jan Seitz, Christian Wittmann, delegates at the 2018 EAA Annual Congress in Milan, at the Ruhr-University of Bochum, Bergische University of Wuppertal and University of Bayreuth. Thanks to Brian Bloch for his editing of the English. We also thank Isabell Keller, Elisabeth Kuhn, Simon Lemnitzer, Marie-Thérèse Meyer, and Anna Mollat for their excellent research assistance.

1 Introduction

The ability of financial statement information (e.g., accruals) to predict future outcomes depends partly on the level of uncertainty in the reporting environment of each firm (Yeung 2009). Uncertainties in financial statements exist to a large extent because of the use of subjective judgements, estimates and assumptions in the measurement process that are inherently unreliable (SEC 2011a). As the level of measurement uncertainty increases, financial statement users find it increasingly more difficult to interpret accounting numbers, such as net income, in order to predict future cash flows (Barth 2006; SEC 2011a). Accordingly, it is necessary to identify precisely which assets and liabilities are affected by measurement uncertainties, to understand their effect on a firm's financial condition and to incorporate them appropriately in one's own investment decisions. For years, standard setters, auditors, and financial statement preparers have constantly been dealing with how to communicate uncertainties in business transactions and accounting estimates, which exert a significant effect on a firm's financial performance (Christensen et al. 2012; Majors 2016; SEC 2011b). Increasing business complexity, dynamic changes in the social, technological, political, and economic environment, as well as the use of highly subjective future-orientated estimation models, have amplified these efforts in recent years (AICPA 1994; Christensen et al. 2012; Christensen et al. 2014; Eilifsen et al. 2017; Mayorga and Sidhu 2012; Lev et al. 2010).

In the early 2000s, the Securities and Exchange Commission (SEC) proposed new disclosure requirements within a firm's Management, Discussion and Analysis (MD&A), in order to enhance investor understanding of measurement uncertainties in a company's financial statements (SEC 2001). Since 2001, firms have been encouraged to disclose their '*critical accounting policies*' (CAPs),[30] which are those accounting policies and estimates that require "management's most difficult, subjective, or complex judgements,

[30] In their initial Financial Reporting Release 60 (FR-60), *Cautionary Advice Regarding Disclosure About Critical Accounting Policies* in 2001 (SEC 2001), the SEC refers to CAPs, whereas in their Proposed Rule, *Disclosure in Management's Discussion and Analysis about the Application of Critical Accounting Policies*, in 2002 (SEC 2002a) and FR-72, *Interpretation: Commission Guidance Regarding Management's Discussion and Analysis of Financial Condition and Results of Operations*, in 2003 (SEC 2003a), they mainly use the term '*critical accounting estimates*' (CAEs). The SEC defines CAEs as those judgmental and subjective estimates involved in the application of CAPs with material impact on a firm's financial condition (SEC 2002a). However, U.S. firms use both terms to designate the corresponding CAP section within their MD&A, and thus do not differentiate adequately between these two

often as a result of the need to make estimates about the effect of matters that are inherently uncertain" (SEC 2001, p. 1). An accounting policy must be flagged as a CAP if changes in the underlying estimates are uncertain, and thus have a material impact on the presentation of a firm's financial condition. Amongst others, economic uncertainty and management judgement are typical sources of such sensitive estimates.

Yet, it remains unclear whether CAPs actually reflect individual uncertain financial statement positions. On the one hand, firms may have an incentive to view CAPs as a mere 'compliance exercise' or as a way to mitigate litigation (e.g., Levine and Smith 2011) and provide standardized boilerplate information without any real information content. On the other hand, if firms use CAPs in the regulators' intended manner (e.g., SEC 2001, 2002a, 2003a), they should be useful for identifying uncertain financial statements (i.e., accruals). Our study provides evidence by examining whether accruals are inherently more uncertain if they are classified as CAPs. Specifically, we examine whether 'uncertain' (if the component is flagged as a CAP) accruals are less persistent with respect to future cash flows than accruals that are 'certain' (if the component is *not* flagged as a CAP). To the best of our knowledge, this is the first study to analyze empirically the persistence of single uncertain accounting accruals that are flagged as CAPs with respect to future cash flows. As a result, we answer a call for additional research (Cole and Jones 2005) and extent prior literature with respect to examining whether CAPs fulfil their intended purpose (Glendening 2017).[31]

According to our results, we provide several relevant new insights. We find that accruals that are flagged as CAPs are less persistent than accruals that are not flagged as CAPs with respect to future cash flows. This is consistent with the SEC's intention that CAPs should capture instances of greater measurement uncertainty embedded in the measurement process of the underlying accruals (e.g., Glendening 2017; SEC 2002a). In an additional analysis, we show that the effect of measurement uncertainties on the persistence of uncertain accruals is not equal across firms and industries. Specifically, we extend prior

categories (Fülbier et al. 2017). We use CAPs and CAEs interchangeably and refer mainly to CAPs within this study.

[31] We do not focus on the usefulness of related CAP disclosures for improving predictions about future earnings or cash flows. The main idea of our study is to analyze whether CAPs are useful for identifying uncertain accounting accruals and thus, whether uncertain accounting accruals are less persistent with respect to a firm's future cash flow.

literature and find that uncertain accrual components are in fact not less useful in predicting future cash flows *per se*; it also depends on their importance and, to a certain extent, the specificity for a given firm.

Our study makes several important contributions. First, prior research on the usefulness of CAPs in identifying uncertain accounting positions is rare. Glendening (2017) finds that the predictive value of earnings is lower when firms provide '*critical accounting estimate*' (CAE[32]) disclosures. Chen et al. (2019) conclude that accruals which require more estimation (based on qualitative information in a firm's footnote and CAP disclosures) are less persistent. However, both studies capture the level of measurement uncertainty on a largely aggregated basis and do not distinguish between the persistence of single accruals that are (not) susceptible to estimation errors. Compared to prior studies, our research design allows for a direct observation of whether CAPs provide information about the persistence of single accrual-based measures. Second, the type of uncertain positions as well as estimates varies substantially between firms, industries, and over time. We use a firm-specific measure that allows controlling for cross-sectional and temporal variation, in order to analyze whether the lower persistence of accruals depends on other factors such as the number of uncertain accruals, their importance or specificity for a given firm and industry. Third, we build upon initial empirical evidence suggesting that investors may benefit from corporate disclosures about uncertain accounting policies and estimates, so as to assess which accruals have more predictive power with respect to predicting future cash flows (Hope 2003a, 2003b; Wolk et al. 2017; Xie 2001). However, there is increasing concern that current disclosure requirements do not serve their intended purpose, because of too much boilerplate information (e.g., Glendening 2017). In this context, we contribute to the debate regarding corporate disclosures about measurement uncertainties in analyzing whether CAPs provide information about the credibility of financial statements. Overall, our findings suggest that CAPs are potentially useful for analysts as well as investors, in enabling them to catch up on measurement uncertainties in financial reports. This allows financial statement users to enhance their understanding of a firm's highly uncertain accounting policies and estimates.

[32] As stated before, we use the term CAP in our study, which comprises both terms introduced by the SEC. However, some studies use mainly the term CAE (e.g., Glendening 2012; Glendening 2017). Thus, in describing their findings, we will use the term CAE to ensure consistency.

We structure our paper as follows. First, we present our theoretical and regulatory framework and review the literature. Second, we develop our hypothesis, present our empirical methods, and describe the data. Third, we exhibit and discuss our results and subject them to a range of robustness checks. The final section concludes.

2 Corporate Disclosures about Measurement Uncertainties

2.1 Relevance of Communicating Measurement Uncertainties

The measurement of financial statements has increasingly been challenged by turbulent and changing market and business environments, as well as an increasing move towards fair value measurement of assets and liabilities (Christensen et al. 2012; Lev et al. 2010). To value financial statement positions accurately, managers have to process information from the market, business, and accounting environments (Palepu et al. 2016). However, fluctuations in such factors as interest and inflation rates, uncertain business models, unforeseen market developments, higher risk industries, as well as accounting distortion due to highly subjective and complex valuation models, lead to serious irreducible measurement uncertainties in applying appropriate accounting policies and estimates (e.g., Palepu et al. 2016). This results in incomplete accounting information as well as an imprecise knowledge of valuation inputs. All this leads to a range of possible outcomes and it is impossible to know which will in fact occur (Beaver 1991; Duffie and Lando 2001; Lu et al. 2010). As a result, managers can neither derive precise probability distributions to forecast future outcomes (Bird et al. 2014) nor make reliable valuations of their assets and liabilities (Beaver 1991). Consequently, it is important to communicate effectively the sources of measurement uncertainty, so that users can incorporate such information into their decisions (Eilifsen et al. 2017).

The usefulness of financial statement information "depends significantly on the user's understanding of the accounting policies followed by the entity" (APB 1972, No. 7). As a result, understanding the measurement basis that shapes the value of financial statement positions is crucial. However, the application of accounting policies, estimates, as well as the selection of valuation inputs depends on internal decisions and information made by managers that, in some parts, is neither available through public channels nor common knowledge. From the perspective of outsiders, it is thus difficult to identify uncertain

positions, because single financial statement items alone cannot provide such information. This increases the relevance of corporate disclosures, because they have the potential to convey proprietary information about managerial estimates and projections (Lev et al. 2010). The literature has also focused on this issue and argues that investors may benefit from disclosures in order to analyze the persistence of financial statement numbers (Xie 2001). Hope (2003b) and Wolk et al. (2017) state that corporate disclosures about sensitive and complex accounting policies and related estimates help outsiders to understand the nature and extent of measurement uncertainties. Consequently, financial statement users may be able to verify the quality of reported financial information (Chartered Professional Accountants Canada 2016; Hope 2003a; Wolk et al. 2017). Further studies conclude that corporate disclosures with respect to measurement uncertainties have the potential to inform outsiders about the (potential) effects on a firm's financial condition if sudden unforeseen changes occur (Campbell et al. 2003; Chartered Professional Accountants Canada 2016; Gietzmann and Trombetta 2003; Healy and Palepu 1993). It is important that firms provide convincing disclosures as to which financial statement positions are inherently more uncertain, so that financial statement users can make 'better' and more informed decisions about their buy, hold, and sell strategy (Barth et al. 2016).

2.2 SEC Regulation on Critical Accounting Policy Disclosures

In the early 2000s, the SEC proposed disclosure requirements within the MD&A to enhance investor understanding of measurement uncertainties in companys' financial statements that are particularly difficult for the management to determine, due to significant subjective judgments and estimations (SEC 2001). Since 2001, firms have been encouraged to provide narrative and quantitative disclosures about their CAPs. The SEC defines them as those accounting policies and estimates that require "management's most difficult, subjective, or complex judgements, often as a result of the need to make estimates about the effect of matters that are inherently uncertain" (SEC 2001, p. 1).

Guidance related to the content of CAP disclosures is included in several SEC releases. In 2001, the SEC issued FR-60 and encouraged firms to include a discussion of the uncertainties embedded in the measurement process of financial statements in their MD&A (SEC 2001). Subsequently, in May 2002, the SEC issued a Proposed Rule regarding CAP disclosures (SEC 2002a). According to the Proposed Rule, firms have to disclose detailed information about the estimate, the methodology, and any highly uncertain assumption

underlying the estimate (SEC 2002a).[33] In December 2003, the SEC specified its requirements in FR-72 and focused even more on those accounting estimates with a significant level of subjectivity and judgement on highly uncertain matters (SEC 2003a). A company should provide disclosures only if the impact of the estimate on a firm's financial condition is material. Based on the interpretive guidance, firms are required to address how they came up with the estimate, how accurate it has been in the past, and how much it has changed, as well as whether it is reasonably likely that the estimate will change in the future. Moreover, each company should analyze the sensitivity of each estimate with respect to a firm's financial performance (SEC 2003a). Although CAP disclosures are already presented separately from the notes, so as to reduce misleading inferences with other financial statement information, the SEC has emphasized that firms should not duplicate accounting policy disclosures already disclosed in the notes. However, no final draft has yet been released, but firms are advised to comply with prior releases of the SEC (e.g., FR-60, Proposed Rule and FR-72).

CAP disclosures should comprise detailed information about highly uncertain estimates, which are in fact ubiquitous in accrual-based accounting and have a significant effect on a firm's (future) financial presentation. In the context of our study, we analyze whether CAPs actually represent highly uncertain financial statement positions. If firms use CAPs in the regulators' intended manner, financial statement users may be able to assess which balance sheet and income statement positions constitute measurement uncertainties. Thus, CAPs have the potential to provide proprietary information, which enhances investor understanding of a firm's highly subjective and complex accounting estimates.

3 Prior Literature

To this day, only a handful of studies focus on whether CAP disclosures provide information about measurement uncertainties in financial statements. Levine and Smith (2011) find that firms with ex-ante higher litigation risk are more likely to provide CAP disclosures, indicating that firms use this disclosure practice to reduce the risk of litigation. The authors further show that the disclosure decision has predictive ability for changes in firm

[33] Furthermore, firms should include both quantitative and qualitative discussions about material changes that may occur, an explanation as to why different estimates would have a material impact on a firm's financial condition and a statement as to whether or not the management has discussed the selection of each CAP with the audit committee.

fundamentals that indicate potential uncertainty surrounding the underlying financial statements. They conclude that firms with fewer (more) CAPs than expected report more (less) reliable reported earnings. Glendening (2012) analyzes whether the presence of a CAE disclosure is associated with the value relevance of financial statement items. He finds that the value relevance of financial positions is negatively associated with the presence of a CAE disclosure and concludes that investors then perceive the underlying balance sheet items as less reliable. In a more recent study, Glendening (2017) examines how the predictive ability of current aggregated earnings with respect to future cash flows varies in the presence of CAE disclosures. He finds a negative association of the predictive value of earnings with respect to future cash flows if there are disclosures about highly uncertain accounting estimates. However, he only focuses on whether there are disclosures about highly uncertain accounting estimates, rather than on which (uncertain) accrual components are (not) useful for predicting future cash flows.

The SEC's intention is to enhance investors understanding of the existence, nature, and impact of uncertain estimates for which management exercises significant managerial discretion. Thus, firms should flag each single accounting policy and/or the underlying accounting topic as uncertain, that fulfil the definition of a CAP. Having said that, CAPs have the potential to provide firm-specific information about measurement uncertainties that are embedded in *single* accrual-based measures. However, the aforementioned studies consider neither which single accrual-based measures are classified as uncertain, nor whether there are accruals that are not susceptible to estimation errors. Thus, current knowledge about the usefulness of CAP disclosures for the identification of uncertain financial statement positions is limited.

4 Hypothesis Development

A fundamental objective of financial reporting is to provide useful information to financial statement users about the amount, timing, and uncertainty of future cash flows (FASB 2010; Glendening 2017). In preparing of financial statements, managers have to apply accounting policies and estimates based on the assessment of present and expected future inflows and outflows associated with their assets and liabilities (Barth et al. 2016). Afterwards, accruals are used to adjust cash flows to reflect their expectations about the future

(Barth 2006; Mayorga and Sidhu 2012). Because accrual accounting incorporates managers' expectations about future cash flows, the Financial Accounting Standards Board (FASB) asserts that such accruals and their components have most predictive power with respect to future cash flow predictions (FASB 1978, 2010). In this context, several studies have confirmed that total accruals (e.g., Dechow 1994; Dechow et al. 1998; Greenberg et al. 1986) as well as disaggregating accruals into their components enhance future cash flow predictions (Barth et al. 2001; Barth et al. 2016). Nevertheless, numerous prior studies find that accruals are less persistent with respect to future cash flows (e.g., Allen et al. 2013; Dechow and Dichev 2002; Dechow and Ge 2006; Richardson et al. 2005, 2006; Xie 2001). More specifically, Sloan (1996) finds that accruals are less persistent due to the subjectivity in their estimation. Lev et al. (2010) argue that accruals based on estimates are less reliable for decision-making purposes. Richardson et al. (2005) demonstrate that the lower persistence of accruals is mostly attributable to those accruals affected by subjectivity and thus measurement uncertainties. More recent studies confirm prior results (e.g., Allen et al. 2013; Christensen et al. 2014). In sum, estimations, complex valuations, and managerial discretion might reduce the benefits of accruals with respect to future cash flow predictions due to objective difficulties and related measurement uncertainties (Glendening 2017; Lev et al. 2010). However, it is therefore important to identify single uncertain accrual components to assess which accruals are (less) useful for predicting future cash flows. In the following, we analyze whether CAP disclosures provide such information.

Analyzing the predictive value of single accounting accruals that are flagged as CAPs is far from straightforward. On the one hand, it is probable that the proposed regulation of the SEC does not fulfil its intended purpose because firms might see CAP disclosures as a mere 'compliance exercise' or as a way to avoid litigation without real information content. Thus, firms may disclose CAPs without taking into account the uncertainty of the underlying accounting position. On the other hand, accounting estimates and managerial projections provide forward-looking proprietary information about the underlying financial statements (Lev et al. 2010). If management correctly identifies uncertainties in their accounting estimates, they might consider such information in the measurement process of financial statements, thus leading not to diminished predictive ability with respect to future cash flows. In contrast, following the literature, measurement uncertainties might result in incomplete accounting information as well as an imprecise knowledge of

valuation inputs. Consequently, management can neither predict future outcomes accurately, nor make reliable valuations, both of which lead to a reduced predictive power of accruals (Dechow and Dichev 2002). Based on the above arguments, we argue that analyzing the predictive value of accruals that are classified as CAPs remains an empirical question.

We consider theoretical and empirical evidence and hypothesize that the disclosure of CAPs (non-disclosure) captures greater (lower) instances of measurement uncertainties embedded in accrual positions. Assuming that CAPs correctly identify highly uncertain accounting positions, 'uncertain' (if the component is flagged as a CAP) accruals should exhibit lower predictive ability rather than accruals that are 'certain' (if the component is *not* flagged as a CAP). Thus, CAPs should induce errors in accrual estimates, thereby limiting the usefulness of accruals to predict future cash flows. As a result, we predict that uncertain accrual components are less persistent with respect to future cash flows and state our first hypothesis (H1) as follows:

***H1**: Accruals that are flagged as CAPs have lower predictive power with respect to future cash flows.*

The type and importance of uncertain positions, as well as of estimates, varies strongly between firms, industries, and over time. First, there are huge discrepancies in the number of uncertain financial statement accounts as well as their importance to a given firm. Therefore, the amount of uncertain positions and the resulting effect on a firm's financial condition may be firm-specific. Based on this assumption, we assume that predictive power depends on the importance to a given firm. Specifically, we argue that predictive power is even lower if the amount of total uncertain CAP accruals reflects a material proportion compared to total accruals ('accrual importance'). Second, there is a wide dispersion in the types of CAPs across firms and industries (e.g., Fülbier et al. 2017; Levine and Smith 2011). While some accounting positions are by nature subject to estimation, others occur rarely in certain companies and industries. On the one hand, firms within the same industry might have equal CAPs because of similar business transactions, business models, and accounting strategies. On the other hand, while there are CAPs in almost every firm and are common within some industries (e.g., deferred taxes, pensions, intangible assets), others are more unusual (e.g., regulatory accounting, leases, investments). We define the former as unspecific and the latter as specific CAPs. If firms have CAPs

that are more unusual compared to their peers, auditors, the audit committee, as well as the management would place greater emphasis on such estimations, because of their uncommon occurrence, leading to estimates that are more precise. In contrast, it might be more challenging for the management to make reliable estimates with respect to the underlying transaction because of a lack of experience in dealing with such estimates.

We argue that accrual importance and accrual specificity might be two moderators that affect the predictive power of uncertain accruals. Consistent with our first hypothesis, we assume that uncertain accruals (those that are flagged as CAPs) have even less predictive power with respect to future cash flows if they are more specific and important to a given firm. Therefore, we state our second hypothesis (H2) as follows:

***H2**: Predictive power is lower for more specific and important accruals that are flagged as CAPs.*

5 Research Design

5.1 Sample Selection

We begin our sample construction by using the S&P 500 composition as of 31.12.2016. We then extract all available 10-Ks from Electronic Data Gathering, Analysis, and Retrieval (EDGAR) during 2002 and 2016.

We use this period because the SEC's first announcement of CAP disclosures was in December 2001 and the majority of firms began providing CAP disclosures in 2002 (cf. Table 12). We obtain accounting data from Worldscope. Our sample selection procedure is as follows. First, while we are interested in the prediction of future cash flows, we limit our sample to those firms with available cash flow data. Second, we eliminate those observations without sufficient financial data for all our analyses. Third, we eliminate all firms without complete time series data. As a result, we remain with 284 firms and 4,260

firm-year observations that span the period 2002 – 2016. Table 11 describes our sample selection procedure:

Table 11
Sample Selection Procedure

		No. of Observations
	Total Number of Observations with available 10-K between 2002 and 2016	7,280
-	Observations without Cash Flow and Sufficient Financial Data	1,811
-	Observations without Complete Time Series Data	1,209
=	Final Sample	4,260

To obtain information about a firm's highly uncertain accounting policies and estimates, we extract the CAP section from the MD&A using Python and collect all CAP headings from each observation. Afterwards, we use a keyword-based coding system to assign each CAP to a single accounting position. Our coding system consists of thirty accounting positions and is based on the FASB Taxonomy and the study of Levine and Smith (2011). This procedure allows us to assess specific financial statement positions that are flagged as highly uncertain, and are thus affected by measurement uncertainties.

5.2 Empirical Framework Analysis 1: Accrual Persistence

5.2.1 Basic Empirical Model

Our initial empirical model for all our analyses is based on the following prediction model (Equation 3). To forecast future cash flows, we follow Barth et al. (2001) and disaggregate earnings into their accrual components. Specifically, the model includes future operating cash flow (CFO_{t+1}), current operating cash flow (CFO_t), changes in accounts receivables (ΔAR), inventories (ΔINV), accounts payable (ΔAP), the total amount of depreciation, amortization and depletion (DEPR), and other accruals (OTHER). Industry (IND) and year (YEAR) dummies are included in all our regressions.

$$CFO_{i,t+1} = \alpha_0 + \beta_1 CFO_{i,t} + \beta_2 \Delta AR_{i,t} + \beta_3 \Delta INV_{i,t} + \beta_4 \Delta AP_{i,t} + \beta_5 DEPR_{i,t} + \beta_6 OTHER_{i,t} + \beta_j \Sigma_j IND_i + \beta_k \Sigma_k YEAR_t + \varepsilon \quad (3)$$

5.2.2 Modeling 'Certain' and 'Uncertain' Accruals

With respect to our first hypothesis, we assume that uncertain accruals (those that are flagged as CAPs) have less predictive power with respect to future cash flows than certain accruals (those that are not flagged as CAPs). To decompose accruals into 'certain' and 'uncertain' accruals, we use a two-step approach that is generally based on the research design of Allen et al. (2013). In a first step, we regress total accruals (ACC) on the disaggregated accrual components of the Barth et al. (2001) model at the industry level. This model takes the following form (Equation 4):

$$ACC_{i,t} = \alpha_0 + \beta_1 \Delta AR_{i,t} + \beta_2 \Delta INV_{i,t} + \beta_3 \Delta AP_{i,t} + \beta_4 DEPR_{i,t} + \beta_5 OTHER_{i,t} + \varepsilon \quad (4)$$

In a second step, we use our hand-collected data on CAP disclosures to identify whether each accrual component of Equation 3 is 'certain' or 'uncertain'. Specifically, we use the coefficients of Equation 4 and aggregate the fitted values of those accruals that are flagged as CAPs as ACC_UNCERTAIN and the fitted values of those accruals that are not flagged as CAPs, as well as the intercept as ACC_CERTAIN. We further allocate the variable OTHER and the error term to ACC_OTHER, because we cannot assign single CAPs to these variables. Finally, we replace the disaggregated earning components of Equation 3 with the values of ACC_CERTAIN, ACC_UNCERTAIN, and ACC_OTHER. Appendix B provides a general example to clarify our procedure. The final model for testing whether 'uncertain' (CAP) accruals have a lower persistence with respect to future cash flows than 'certain' (non-CAP) accruals takes the following form (Equation 5):[34]

$$CFO_{i,t+1} = \alpha_0 + \beta_1 CFO_{i,t} + \beta_2 ACC\_CERTAIN_{i,t} + \beta_3 ACC\_UNCERTAIN_{i,t} + \beta_4 ACC\_OTHER_{i,t} + \beta_j \Sigma_j IND_i + \beta_k \Sigma_k YEAR_t + \varepsilon \quad (5)$$

[34] As a robustness check, we also include the sum of cash flows from t+1 to t+3 as our main dependent variable to control for the long term realization of cash flows from accruals.

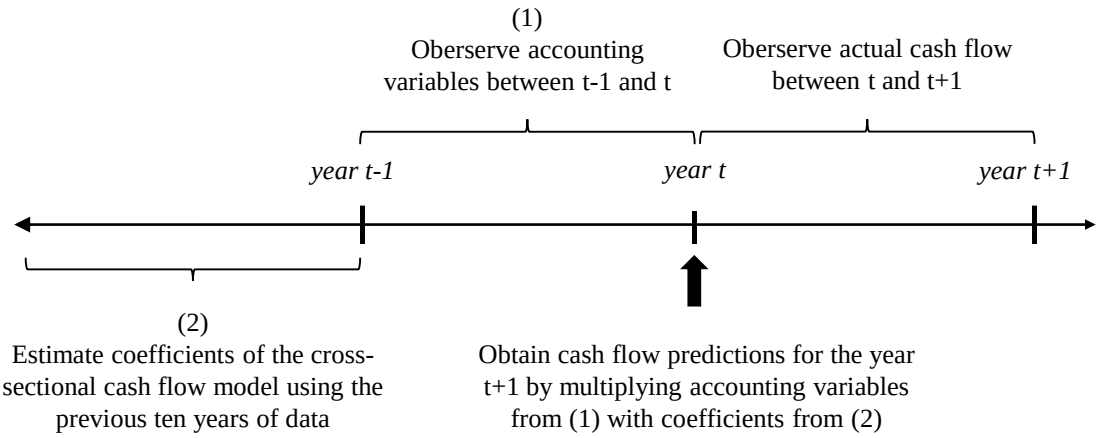
5.3 Empirical Framework Analysis 2: Important and Specific Accruals

5.3.1 Cross-Sectional Cash Flow Prediction Model

With respect to our second hypothesis, we analyze whether the predictive power of uncertain accruals is determined by their importance and/or specificity. We test the usefulness of important and specific uncertain accruals in terms of their ability to forecast future cash flows. Assuming that important and specific uncertain accruals are less useful, their ability to predict future cash flows should decline, resulting in an increased prediction error. Our research design for our second hypothesis consists of two stages. In the first, we estimate firm-specific cash flow forecasts up to one year. A simple regression of a variable on lagged values of the same variable is not a test of predictive ability (Lev et al. 2010). In this context, Poon and Granger (2003) argue that a good forecast model should "withstand the robustness of an out-of-sample test, a test design that is closer to reality" (Poon and Granger 2003, p. 492). Therefore, we use in-sample and out-of-sample regressions to forecast the next period's cash flows. We follow the general procedure of Hou et al. (2012) and Lev et al. (2010) by using in-sample tests of cash flows regressed on lagged values of the independent variables from Equation 3, using the previous ten years of data. By obtaining coefficients of the in-sample forecast, we estimate out-of-sample predictions of cash flows for our final sample, using the model of Barth et al. (2001). Finally, we determine the prediction error for each firm by subtracting the predicted from the current cash flow. Figure 6 presents a timeline for the estimation procedure described above (Hou et al. 2012). Appendix C provides a general example to clarify our prediction procedure (example to forecast 2002 operating cash flow).[35]

[35] Although our final sample covers the period 2002 – 2016, we add additional financial variables from 2000, 2001, and 2017 to run all our analysis. By doing so, we do not lose any observations.

Figure 6
Timeline of Cash Flow Forecasts



To evaluate the effect of measurement uncertainties contained in accounting positions and thus, whether or not they are flagged as CAPs, we conduct univariate statistical tests and analyze the pooled firm-specific mean absolute prediction error (MAPE) of our prediction model.

5.3.2 Measurement of Importance and Specificity

Our second hypothesis posits that the level of measurement uncertainty depends on the importance and specificity of the underlying uncertain financial statement accounts. We argue that the predictive power of accruals will be lower for more specific (SPECIFICITY) and more important (IMPORTANCE) uncertain accruals. We define IMPORTANCE as the sum of all uncertain accrual components that are classified as CAPs, scaled by the sum of all accrual components of Equation 4. We argue that the predictive power of uncertain accruals with respect to future cash flows might be lower if IMPORTANCE is relatively high. Thus, we check whether accrual size affects our findings. Moreover, the effect of uncertain accrual components on the predictability of future cash flows may also depend upon their specificity. Using the proportion of single uncertain accrual components and the total number of uncertain accrual components may solve this problem. However, we argue that the effect depends on the occurrence and thus, on the specificity compared to their peers.

Our measure of SPECIFICITY (Equation 6) is based on a common term-weighting scheme from the textual analysis literature, namely term frequency-inverse document frequency (*tf-idf*) (Loughran and McDonald 2011).[36] Applied to our study, the general form of SPECIFICITY is then:

$$SPECIFICITY = \begin{cases} \frac{(1+log(cap\_i))}{(1+log(i))} \times log\frac{N}{cap\_n} & if\ cap\_i \geq 1 \\ 0 & otherwise \end{cases} \quad (6)$$

SPECIFICITY is estimated for each accrual component separately.[37] We define Equation 6 as accrual specificity with *cap_i* (*cap_n*) as the raw count of the number of firms in a given year and industry (raw count of the number of firms in a given year), that classify an accrual component of Equation 4 as critical. *N* is the total number of firms in a given year and *i* the total number of firms per year and industry. SPECIFICTY is the mean of all four specificity scores. To test our second hypothesis, we split our sample by the median of IMPORTANCE and SPECIFICTY and compare the mean absolute prediction error across these four groups.

6 Results

6.1 Analysis 1: Accrual Persistence

6.1.1 Descriptive Statistics

Table 12 provides descriptive statistics about the number of firms providing CAP disclosures, as well as the classification of each accrual component of Equation 3 as critical. Only four firms in 2002 and one in 2003 do not provide any CAP disclosures. In the following years, all firms in our sample provide a CAP discussion in their MD&A section. About 90 percent of our sample classify their depreciation, amortization and depletion as highly uncertain.[38] Compared to that, less than 10 percent of the sample have uncertain

[36] In their study, Loughran and McDonald (2011) find that using *tf-idf* leads to better results than using simple proportions.

[37] Thus, we estimate SPECIFICTY_DEPR for the specificity of depreciation, amortization and depletion, SEPCIFICTY_AR for accounts receivables, SPECIFITICY_AP for accounts payables, and SPECIFICTY_INV for inventories.

[38] DEPR includes those CAPs that relate to accounting topics such as property, plant, and equipment, as well as intangibles.

accounts payable. Moreover, around 30 percent (35 %) of the sample have accounts receivables (inventories) that are highly uncertain.

Table 12
Descriptive Statistics of CAP Disclosures

Year	CAP=1	CAP=0	AR	%	INV	%	AP	%	DEPR	%	N
2002	280	4	89	31.79	99	35.36	24	8.57	204	72.86	284
2003	283	1	95	33.57	107	37.81	21	7.42	224	79.15	284
2004	284	0	98	34.51	102	35.92	24	8.45	232	81.69	284
2005	284	0	95	33.45	103	36.27	21	7.39	231	81.34	284
2006	284	0	90	31.69	101	35.56	23	8.10	233	82.04	284
2007	284	0	84	29.58	102	35.92	23	8.10	243	85.56	284
2008	284	0	86	30.28	103	36.27	23	8.10	249	87.68	284
2009	284	0	85	29.93	103	36.27	21	7.39	251	88.38	284
2010	284	0	81	28.52	103	36.27	23	8.10	251	88.38	284
2011	284	0	79	27.82	101	35.56	24	8.45	250	88.03	284
2012	284	0	76	26.76	98	34.51	26	9.15	251	88.38	284
2013	284	0	75	26.41	98	34.51	27	9.51	250	88.03	284
2014	284	0	69	24.30	94	33.10	26	9.15	250	88.03	284
2015	284	0	66	23.24	99	34.86	27	9.51	252	88.73	284
2016	284	0	62	21.83	97	34.15	25	8.80	253	89.08	284

Table 12 presents summary statistics about the number of firms that classified each accrual component of the Barth et al. (2001) model as a CAP. AR is accounts receivables, INV is inventories, AP is accounts payable and DEPR is depreciation, amortization and depletion. See Appendix A for detailed variable descriptions.

Table 13 presents summary statistics for the main variables of the Cash Flow Prediction Model (Equation 3), and the Certain and Uncertain Accrual Model (Equation 5). Current cash flows has a mean value (median) of 13.4 percent (12.1 %). Consistent with prior studies (e.g., Barth et al. 2001; Sloan 1996), we find that the means and medians of CFO are positive, and those of aggregated accruals (ACC = EARN – CFO), as well as ACC_CERTAIN, ACC_UNCERTAIN, and ACC_OTHER are negative. This is because aggregated accruals include depreciation, amortization and depletion, but the acquisition of depreciable and amortizable assets is related to investing, and not to the operating cash flow (Barth et al. 2001; El-Sayed Ebaid 2011). Moreover, the magnitudes of ΔAR, ΔINV and ΔAP are smaller compared to DEPR. On average, firms disclose six CAPs. The absolute number of CAPs ranges between zero and 14 CAPs. This finding is in line with the assumption that the number and types of CAPs vary between firms. Moreover, the number of uncertain accrual components (CAP_ACCRUAL) ranges from zero to four with a

mean of two (median of two). About 30 percent of all CAPs relate to those accruals included in Equation 4. The number of all other CAPs (CAP_OTHER) ranges from zero to ten with a mean (median) of four.

Table 13
Descriptive Statistics Financial Variables Hypothesis 1

Variables	Mean	Min	25 %	Median	75 %	Max	SD	N
CFO	0.134	-0.064	0.081	0.121	0.172	0.423	0.079	4,260
ΔAR	-0.002	-0.182	-0.011	0.000	0.010	0.129	0.036	4,260
ΔINV	-0.002	-0.133	-0.006	0.000	0.004	0.091	0.026	4,260
ΔAP	-0.001	-0.122	-0.007	0.000	0.007	0.091	0.024	4,260
DEPR	0.044	0.003	0.028	0.039	0.053	0.140	0.025	4,260
OTHER	-0.009	-0.254	-0.030	-0.008	0.012	0.237	0.060	4,260
ACC	-0.053	-0.292	-0.072	-0.046	-0.025	0.095	0.052	4,260
ACC_CERTAIN	-0.014	-0.294	-0.023	-0.009	0.001	0.196	0.035	4,260
ACC_UNCERTAIN	-0.032	-0.330	-0.046	-0.028	-0.012	0.178	0.040	4,260
ACC_OTHER	-0.006	-0.376	-0.022	-0.005	0.012	0.210	0.044	4,260
CAP_TOTAL	6	0	5	6	7	14	3	4,260
CAP_ACCRUAL	2	0	1	2	2	4	1	4,260
CAP_OTHER	4	0	4	4	5	10	2	4,260

Table 13 presents summary descriptive statistics for all firm-year observations for the period 2002 through 2016. CFO is operating cash flow. ΔAR is the year-to-year change in accounts receivables. ΔINV is the year-to-year change in inventories. ΔAP is the year-to-year change in accounts payable. DEPR is depreciation, amortization and depletion. OTHER is the difference between total accruals and ΔAR, ΔINV, ΔAP and DEPR. ACC is earnings minus operating cash flows. ACC_CERTAIN are the fitted values from those accruals that are not flagged as CAPs and the intercept. ACC_UNCERTAIN are the fitted values from those accruals that are flagged as CAPs. ACC_OTHER are the fitted values from all other accruals and the residuals from Equation 4. CAP_TOTAL is the total number of CAPs. CAP_ACCRUAL is the number of accrual components of the Barth et al. (2001) model that are classified as CAPs. CAP_OTHER is the number of all other CAPs that are not already included in CAP_ACCRUAL. See Appendix A for detailed variable descriptions.

Because (multi-)collinearity may be a problem, we present a pairwise Pearson correlation matrix in Table 14 and Variance Inflation Factors (VIFs) in Table 15. As expected, ACC is significantly negatively correlated with CFO. With the exception of ΔINV, all accrual components are significantly correlated with CFO and are generally correlated significantly with each other. These findings are in line with the literature. The correlations of ACC_UNCERTAIN, ACC_CERTAIN and ACC_OTHER with CFO are negative. If accrual components relate to ACC_CERTAIN, they do not relate simultaneously to ACC_UNCERTAIN, ACC_OTHER respectively. In addition, the highest VIF is 2.23 for ΔAR and 2.12 for ΔAP. Thus, all correlations as well as VIFs are below established critical levels (Wooldridge 2013), so that we assume (multi-)collinearity not to affect our results.

Table 14
Pearson Correlations

	Variable	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
(1)	CFO	1.00									
(2)	ΔAR	**0.03**	1.00								
(3)	ΔINV	0.00	**0.51**	1.00							
(4)	ΔAP	**0.08**	**0.66**	**0.55**	1.00						
(5)	DEPR	**0.34**	**0.10**	**0.08**	**0.07**	1.00					
(6)	OTHER	**-0.22**	**-0.48**	**-0.47**	**-0.26**	**-0.13**	1.00				
(7)	ACC	**-0.31**	**0.14**	**0.13**	0.02	**-0.47**	**0.54**	1.00			
(8)	ACC_CERTAIN	**-0.10**	**0.35**	**0.22**	0.01	**-0.14**	**-0.23**	**0.18**	1.00		
(9)	ACC_UNCERTAIN	**-0.11**	**0.44**	**0.52**	**0.33**	**-0.32**	**-0.29**	**0.28**	**-0.34**	1.00	
(10)	ACC_OTHER	**-0.14**	**-0.42**	**-0.40**	**-0.22**	**-0.12**	**0.88**	**0.67**	**-0.21**	**-0.24**	1.00

Table 14 presents pairwise Pearson correlations. CFO is operating cash flow. ΔAR is the year-to-year change in accounts receivables. ΔINV is the year-to-year change in inventories. ΔAP is the year-to-year change in accounts payable. DEPR is depreciation, amortization and depletion. OTHER is the difference between total accruals and ΔAR, ΔINV, ΔAP and DEPR. ACC is earnings minus operating cash flows. ACC_UNCERTAIN are the fitted values from those accruals that are flagged as CAPs. ACC_CERTAIN are the fitted values from those accruals that are not flagged as CAPs and the intercept. ACC_OTHER are the fitted values from all other accruals and the residuals from Equation 4. See Appendix A for detailed variable descriptions. Bold indicates significances at the two-tailed 10 % level or higher.

Table 15
Variance Inflation Factors

	Model 1	Model 2
	Cash Flow Prediction Model	Certain and Uncertain Accrual Model
CFO	1.23	1.09
ΔAR	2.23	
ΔINV	1.76	
ΔAP	2.12	
DEPR	1.14	
OTHER	1.63	
ACC_CERTAIN		1.33
ACC_UNCERTAIN		1.36
ACC_OTHER		1.26

Table 15 presents variance inflation factors (VIFs) separately for the Cash Flow Prediction Model (Equation 3) (Model 1) and the Certain and Uncertain Accrual Model (Equation 5) (Model 2) of our main analyses. Dependent variable is CFO in t+1. CFO is operating cash flow. ΔAR is the year-to-year change in accounts receivables. ΔINV is the year-to-year change in inventories. ΔAP is the year-to-year change in accounts payable. DEPR is depreciation, amortization and depletion. OTHER is the difference between total accruals and ΔAR, ΔINV, ΔAP and DEPR. ACC_UNCERTAIN are the fitted values from those accruals that are flagged as CAPs. ACC_CERTAIN are the fitted values from those accruals that are not flagged as CAPs and the intercept. ACC_OTHER are the fitted values from all other accruals and the residuals from Equation 4. See Appendix A for detailed variable descriptions.

6.1.2 Persistence of 'Certain' and 'Uncertain' Accrual Components

First, we turn to our first hypothesis (H1). H1 posits that CAPs are informative with respect to measurement uncertainties within accrual-based measures. Therefore, we predict that accrual components that are flagged as CAPs are less persistent with respect to future cash flows. Our test proceeds in two steps. In the first, we replicate the Barth et al. (2001) model finding that current cash flow disaggregated accrual components have predictive power with respect to future cash flows.

Regression results are reported in Table 16. The first regression (Model 1) is the base regression of current cash flows and disaggregated earnings on future cash flows. Consistent with Barth et al. (2001), we find that ΔAR, ΔINV, DEPR, and OTHER are significantly positively related, whereas ΔAP is significantly negatively related to future cash flows. Moreover, the coefficients of all accrual components are lower than the coefficient of CFO. This corroborates prior results that the accrual components are less persistent than the cash flow component (e.g., Allen et al. 2013; Richardson et al. 2005, 2006; Sloan 1996). In a second step, we use our hand-collected data about CAPs and decompose accruals into ACC_CERTAIN, ACC_UNCERTAIN. We include further ACC_OTHER that captures all other components of Equation 4 that cannot be assigned to single CAPs. This step allows us to test whether uncertain accruals (those that are flagged as CAPs) are less persistent than certain accruals (those that are not flagged as CAPs). With respect to all other accruals (ACC_OTHER), we do not predict the height of the coefficient, because we cannot determine whether the underlying accrual components are uncertain or certain.

Table 16
Regression Results Accrual Persistence

	Model 1		Model 2	
	Cash Flow Prediction Model		Certain and Uncertain Accrual Model	
CFO	0.760	***	0.740	***
	(71.01)		(71.46)	
ΔAR	0.354	***		
	(12.04)			
ΔINV	0.245	***		
	(6.86)			
ΔAP	-0.526	***		
	(-12.46)			
DEPR	0.082	**		
	(2.42)			
OTHER	0.198	***		
	(12.98)			
ACC_CERTAIN			0.141	***
			(5.67)	
ACC_UNCERTAIN			0.066	***
			(3.00)	
ACC_OTHER			0.087	***
			(4.57)	
Constant	0.034	***	0.042	***
	(8.31)		(10.23)	
Fixed Effects	Y,I		Y,I	
R^2	0.647		0.627	
N	4,260		4,260	

Table 16 presents regression results of the Cash Flow Prediction Model (Equation 3) (Model 1) and the Certain and Uncertain Accrual Model (Equation 5) (Model 2). Dependent variable is CFO in t+1. CFO is operating cash flow in t. ΔAR is the year-to-year change in accounts receivables. ΔINV is the year-to-year change in inventories. ΔAP is the year-to-year change in accounts payable. DEPR is depreciation, amortization and depletion. OTHER is the difference between total accruals and ΔAR, ΔINV, ΔAP and DEPR. ACC is earnings minus operating cash flows. ACC_UNCERTAIN are the fitted values from those accruals that are flagged as CAPs. ACC_CERTAIN are the fitted values from those accruals that are not flagged as CAPs and the intercept. ACC_OTHER are the fitted values from all other accruals and the residuals from Equation 4. See Appendix A for detailed variable descriptions. Each model is estimated with year- (Y) and industry- (I) fixed effects. ***, ** and * denotes significance at 1 %, 5 % and 10 % level, respectively.

Model 2 of Table 16 includes ACC_CERTAIN, ACC_UNCERTAIN, and ACC_OTHER instead of each accrual component of the Barth et al. (2001) model (Equation 3). We predict that the coefficient of ACC_UNCERTAIN is lower than the coefficients of CFO and ACC_CERTAIN. The results presented in Table 16 are consistent with this prediction. The coefficient of ACC_UNCERTAIN is 0.066, while the coefficients of CFO and

ACC_CERTAIN are 0.740 and 0.141, respectively. All differences between the coefficients are statistically significant (p-value < 0.01). What seems interesting is that the coefficient of ACC_OTHER is higher than the coefficient of ACC_UNCERTAIN. Thus, we assume that all accruals included in ACC_OTHER are unaffected by (a high degree of) measurement uncertainty. To sum up, we conclude that accruals classified as CAPs being the least persistent components indicating that CAPs depict subjective and uncertain accrual positions.

6.2 Analysis 2: Important and Specific Accruals

6.2.1 Descriptive Statistics

Based on the findings relating to our first hypothesis, we now turn to our second hypothesis (H2). Panel A of Table 17 presents summary statistics of all variables used in the cross-sectional prediction model between 1991 and 2014. Panel B of Table 17 reports the average coefficients from the pooled regressions estimated each year between 2000 and 2014. Panel C of Table 17 presents the observed accounting numbers between 2001 and 2015 that are used to determine predicted cash flows. In all our yearly regressions, current cash flows and all disaggregated accrual components are highly persistent with respect to future cash flows. Consistent with our prior findings, ΔAR, ΔINV, DEPR, and OTHER are significantly positively related to future cash flows, whereas the coefficient of ΔAP is negative and significant in all our yearly regressions.

We predict a firm's operating cash flow in t+1 by multiplying the coefficient in t-1 (Table 17, Panel B) with the corresponding accounting numbers in t (Table 17, Panel C). Panel D of Table 17 presents estimated predicted cash flows (PRED_CFO) between 2002 and 2014, MAPE, estimated by the absolute difference between the actual and predicted cash flow and the mean absolute error term (MAET). The average (median) predicted cash flow is about 11.2 percent (10.4 %). This leads to a pooled mean absolute prediction error of 3.8 percent with a median of 2.6 percent. Comparing the mean absolute prediction error with the pooled mean signed error of the cash flow model (Table 17, Panel D), we conclude that the cross-sectional cash flow prediction model provides satisfactory results.

Table 17

Descriptive Statistics Financial Variables Hypothesis 2

Panel A: Cross-Section Prediction Model (Period: 1991 – 2014)								
Variables	**Mean**	**Min**	**25 %**	**Median**	**75 %**	**Max**	**SD**	**N**
CFO	0.088	-0.341	0.035	0.088	0.146	0.450	0.088	73,199
ΔAR	-0.008	-0.400	-0.028	-0.001	0.022	0.280	-0.008	73,199
ΔINV	-0.004	-0.264	-0.013	0.000	0.009	0.203	-0.004	73,199
ΔAP	-0.003	-0.229	-0.016	-0.000	0.014	0.177	-0.003	73,199
DEPR	0.054	0.005	0.031	0.046	0.066	0.203	0.035	73,199
OTHER	0.001	-0.444	-0.043	-0.003	0.041	0.509	0.126	73,199
Panel B: Average Coefficients from the Pooled Regressions (Period: 2000 – 2014)								
coeff_CFO	0.022	0.019	0.022	0.022	0.023	0.026	0.002	4,260
coeff_ΔAR	0.620	0.589	0.607	0.620	0.638	0.648	0.018	4,260
coeff_ΔINV	0.300	0.291	0.298	0.301	0.302	0.308	0.004	4,260
coeff_ΔAP	0.246	0.213	0.218	0.243	0.269	0.301	0.029	4,260
coeff_DEPR	-0.442	-0.478	-0.466	-0.431	-0.426	-0.418	0.020	4,260
coeff_OTHER	0.191	0.158	0.173	0.183	0.215	0.236	0.022	4,260
coeff_constant	0.167	0.160	0.163	0.166	0.168	0.176	0.004	4,260
Panel C: Observed Accounting Numbers (Period: 2001 – 2015)								
acc_CFO	0.134	-0.064	0.080	0.121	0.174	0.423	0.081	4,260
acc_ΔAR	-0.004	-0.182	-0.013	-0.001	0.010	0.129	0.039	4,260
acc_ΔINV	-0.002	-0.133	-0.007	0.001	0.004	0.091	0.028	4,260
acc_ΔAP	-0.002	-0.122	-0.008	-0.001	0.007	0.091	0.026	4,260
acc_DEPR	0.045	0.003	0.029	0.039	0.054	0.140	0.025	4,260
acc_OTHER	-0.008	-0.254	-0.030	-0.008	0.014	0.237	0.062	4,260
Panel D: Estimated Predicted Cash Flows and Prediction Errors (Period: 2002 – 2016)								
PRED_CFO	0.112	-0.073	0.079	0.104	0.138	0.341	0.050	4,260
MAPE (CFO)	0.038	0.001	0.011	0.026	0.051	0.412	0.040	4,260
MAET (CFO)	0.031	0.001	0.009	0.021	0.040	0.408	0.034	4,260

Table 17 presents summary statistics of the variables used in the in-sample and out-of-sample regressions. Panel A presents summary statistics of the variables used in the cross-sectional prediction model between 1991 and 2014. Panel B reports the average coefficients from the pooled regressions estimated each year between 2000 and 2014 using the previous ten years of data. Panel C presents the observed accounting numbers between 2001 and 2015 that are used to determine predicted cash flows. Panel D presents estimated predicted cash flows (PRED_CFO) between 2002 and 2016 and the mean absolute prediction errors (MAPE). MAET is the mean absolute error term estimated by the absolute value of the residuals of Equation 3. CFO is operating cash flow. ΔAR is the year-to-year change in accounts receivables. ΔINV is the year-to-year change in inventories. ΔAP is the year-to-year change in accounts payable. DEPR is depreciation, amortization and depletion. OTHER is the difference between total accruals and ΔAR, ΔINV, ΔAP and DEPR. See Appendix A for detailed variable descriptions.

6.2.2 Importance and Specificity

Based on our findings for H1, we argue that a (lower) predictive power of uncertain accrual components might also depend on their importance and specificity for a given firm. We hypothesize that the predictive power of uncertain accrual components is even lower for those that are important and specific. Compared to our empirical framework for H1, we analyze the usefulness of important and specific uncertain accruals in terms of their ability to forecast future cash flows. We assume that the absolute prediction error with respect to future cash flow is higher if a firm has more important and more specific uncertain accruals.

Table 18
Descriptive Statistics – Importance and Specificity

Variables	Mean	Min	25 %	Median	75 %	Max	SD	N
SPECIFICITY_AR	0.979	0.000	0.819	1.069	1.161	1.285	0.243	4,260
SPECIFICITY_INV	0.810	0.000	0.699	0.993	1.090	1.271	0.416	4,260
SPECIFICITY_AP	1.009	0.000	0.888	1.087	1.292	1.626	0.394	4,260
SPECIFICITY_DEPR	0.203	0.095	0.143	0.174	0.264	0.423	0.081	4,260
SPECIFICITY	1.885	0.000	1.173	1.707	2.922	4.113	1.056	4,260
IMPORTANCE	0.422	0.001	0.149	0.383	0.685	1.000	0.301	4,260

Table 18 presents summary descriptive statistics of the variables IMPORTANCE and SPECIFICITY. IMPORTANCE is the sum of all uncertain accruals scaled by total accruals. SPECIFICITY is a measure for the relative occurrence of a specific accrual CAP in the firm's respective industry (determined by a procedure similar to *tf-idf* (Loughran and McDonald 2011). It is measured as the mean of SPECIFICITY_AR, SPECIFICITY_INV, SPECIFICITY_AP and SPECIFICITY_DEPR. SPECIFICITY_AR relates to the relative occurrence of classifying accounts receivables as 'critical'. SPECIFICITY_INV relates to the relative occurrence of classifying inventories as 'critical'. SPECIFICITY_AP relates to the relative occurrence of classifying accounts payables as 'critical'. SPECIFICITY_DEPR relates to the relative occurrence of classifying depreciation, amortization and depletion as 'critical'. See Appendix A for detailed variable descriptions.

First, Table 18 presents summary statistics for our two measures IMPORTANCE and SPECIFICITY. As can be seen, the importance of uncertain accruals, estimated by the ratio between the sum of all uncertain accruals and the sum of total accruals, captures on average 42.2 percent with a median of 38.3 percent of all accrual components. By looking at our specificity values, DEPR has the lowest value at 0.203, indicating that the majority of firms classify their depreciation, amortization and depletion as uncertain. Thus, DEPR is the most unspecific uncertain accrual component in our sample. Compared to this, AR and AP are the most specific accrual components with a value of 0.979 and 1.009 respectively. This indicates that a lower number of firms have receivables and/or accounts payable that are affected by measurement uncertainties.

Table 19
Univariate Results – Importance and Specificity

Panel A: Mean Absolute Prediction Error (MAPE)

		SPECIFICITY (median split)		
		Low	High	Diff
IMPORTANCE (median split)	Low	0.030 n = 761	0.035 n = 1,369	0.005 ***
	High	0.042 n = 1,362	0.044 n = 768	0.002
	Diff	0.012 ***	0.009 ***	

Panel B: Mean Absolute Error Term (MAET)

		SPECIFICITY (median split)		
		Low	High	Diff
IMPORTANCE (median split)	Low	0.026 n = 761	0.029 n = 1,369	0.003 **
	High	0.034 n = 1,362	0.032 n = 768	-0.002
	Diff	0.008 ***	0.003 ***	

Table 19 presents results from two-sided t-tests for differences in means of the mean absolute prediction error (MAPE) (Panel A) and mean absolute error term (MAET) (Panel B). The sample (n = 4,260) is splitted by the median of IMPORTANCE and SPECIFICITY into 'low'- (below median) and 'high'-affected (above median) groups, respectively. MAPE is calculated as the difference of current operating cash flow (CFO) and predicted operating cash flow (PRED_CFO). MAET is calculated as the mean absolute error term estimated by the absolute value of the residuals of Equation 3. Presented are the means of these four groups. See Appendix A for detailed variable descriptions. ***, ** and * denotes significance at 1 %, 5 % and 10 % level, respectively.

Second, we split our sample into four groups by the median of IMPORTANCE and SPECIFICITY. Our 2x2 matrix is presented in Table 19. In Panel A, Group 1 (Group 2) includes the MAPE of those observations with less important and less (more) specific uncertain accruals. Group 3 (Group 4) depicts the MAPE of those observations with more important and less (more) specific uncertain accruals. By comparing the absolute prediction error between these four groups, we find that firms in Group 4 have the highest, and firms in Group 1 the lowest prediction error. It is worth noting that the absolute prediction error based on important uncertain accruals (Group 3) is about 1.2 percentage points higher than for unimportant uncertain accruals (Group 1). This result is consistent with our prediction that the predictive value is even lower for more specific and important uncertain accruals, thus indicating a lower predictive ability. Moreover, the MAPE of Group 2 (Group 3) is significantly higher than that of Group 1. It seems that firms with more (less) important (and more specific) uncertain accruals have significantly higher

prediction errors. However, our results seem to be driven primarily by the importance of uncertain accruals, because the results for specificity are marginal and inconclusive when comparing the mean absolute prediction error of firms in Group 3 with Group 4. Even if the difference of MAPE between Group 4 and Group 3 is positive, it remains insignificant. The results for MAET are presented in Panel B and confirm our results.

In summary, our results indicate that the effect of all uncertain accruals on future firm fundamentals is not the same across firms. Furthermore, accrual components classified as CAPs do not have less predictive power with respect to future cash flows *per se*; it also depends mainly on their importance and to some extent on the specificity for a given firm.

7 Robustness Checks

We conduct several sensitivity tests to assess the validity of our results. The first two robustness checks refer to our first hypothesis, whereas the other robustness checks refer to our second hypothesis.

In our first set of robustness checks, we follow prior studies (e.g., Barth et al. 2016; Glendening 2017) and argue that accruals not only affect a firm's next period cash flow, but also cash flows in multiple (future) periods. We employ one alternative test to analyze the persistence of uncertain accrual components with respect to cash flows across multiple periods. We include CFO3Y, which equals the sum of CFO from t+1 to t+3 as our dependent variable in Equation 5, and then re-perform our analysis. The regression results are presented in Table 20. The coefficient of ACC_UNCERTAIN is still lower than that of ACC_CERTAIN and CFO, but remains highly insignificant. Whereas uncertain accruals might still have predictive value for the next period's cash flow, this finding implies that there is no incremental predictive value for cash flows across multiple periods. Thus, we confirm our results that CAPs convey information about imprecise and subjective estimates, thus reducing the predictive value of accruals with respect to future cash flows.

Table 20
Robustness Check I – Long Term Uncertain Accruals

	Certain and Uncertain Accrual Model	
CFO	1.973	***
	(63.02)	
ACC_CERTAIN	0.269	***
	(3.55)	
ACC_UNCERTAIN	0.066	
	(1.01)	
ACC_OTHER	0.147	**
	(2.49)	
Constant	0.183	***
	(15.10)	
Fixed Effects	Y,I	
R^2	0.639	
N	3,408	

Table 20 presents regression results of the Certain and Uncertain Accrual Model (Equation 5) (Robustness Check I). Dependent variable is CFO3Y. CFO3Y equals the sum of CFO from t+1 to t+3. CFO is operating cash flow in t. ACC_UNCERTAIN are the fitted values from those accruals that are flagged as CAPs. ACC_CERTAIN are the fitted values from those accruals that are not flagged as CAPs and the intercept. ACC_OTHER are the fitted values from all other accruals and the residuals from Equation 4. See Appendix A for detailed variable descriptions. The model is estimated with year- (Y) and industry- (I) fixed effects. ***, ** and * denotes significance at 1 %, 5 % and 10 % level, respectively.

Second, the Barth et al. (2001) model that we used is based on four major accrual components. Nevertheless, there may be additional accruals that are also related to future cash flows, but are aggregated in the variable OTHER. Barth et al. (2016) develop the Barth et al. (2001) model further and argue that the role of accruals in predicting future cash flows depends mainly upon their origin, i.e. whether the association of cash flows and accruals has occurred or will occur. Therefore, we follow Barth et al. (2016) and include two major types of accruals for which the associated cash flow occurs in the period after the economic event (e.g., pensions and accounts receivables), as well as before the economic event (e.g., deferred revenue and inventories) (Barth et al. 2016). In sum, they assign 17 different accrual components to both variables.[39] Thus, the final model includes

[39] SFP^A is the sum of total receivables, deferred tax assets minus the sum of accounts payable, accrued expenses, pension liability, income taxes payable, and deferred tax liability. SFP^B is the sum of inventories, prepaid expenses, income tax refunds, property, plant and equipment, intangible assets, deferred charges, investments, advances-equity, long-term pension assets minus deferred revenue (Barth et al. 2016).

a much larger number of accruals and may therefore be more suitable analyzing whether CAP disclosures are useful for determining the persistence of single accounting accruals. We further use CAP disclosures to flag each accrual component included in SFP^A and SFP^B as either 'certain' or 'uncertain'. Using the procedure of our main analysis, we include SFP^A_UNCERTAIN, SFP^A_CERTAIN, SFP^B_UNCERTAIN, and SFP^B_CERTAIN in our final model and rerun our entire analysis.[40] The results with respect to the analysis of our first (second) hypothesis are presented in Table 21 (Table 22). Table 21 reveals that both uncertain components are less persistent than the respective certain components with respect to future cash flows. Interestingly, both coefficients of SFP^B are much lower than SFP^A. This may be because SFP^B contains more long-term accruals (such as intangible assets or property, plant and equipment) which, in particular, do not simply align in one-year ahead cash flows. However, our findings support all inferences by considering a larger number and distinct types of uncertain accrual components in our main analysis. With respect to our second hypothesis (Table 22), we find that the predictive power of more important accruals that are classified as CAPs is slightly higher than for less important accruals. We further find that firms have significantly higher prediction errors if their uncertain accruals are more specific (but less important). Thus, we confirm our main results and conclude that IMPORTANCE and SPECIFICITY are two moderators affecting the forecast ability of uncertain accrual components with respect to future cash flows.

[40] See Appendix D for detailed explanations.

Table 21

Robustness Check II – Accrual Persistence using the Barth et al. (2016) Model

	Model 1		Model 2	
	Cash Flow Prediction Model		Certain and Uncertain Accrual Model	
CFO	0.743	***	0.728	***
	(67.52)		(66.71)	
ΔSFP^A	0.123	***		
	(8.78)			
ΔSFP^B	0.089	***		
	(7.08)			
ΔSFP^A_CERTAIN			0.060	***
			(4.27)	
ΔSFP^A_UNCERTAIN			0.031	***
			(4.58)	
ΔSFP^B_CERTAIN			0.038	***
			(3.09)	
ΔSFP^B_UNCERTAIN			0.029	**
			(2.28)	
OACC	0.093	***	0.034	***
	(6.95)		(2.83)	
Constant	0.051	***	0.051	***
	(12.07)		(11.96)	
Fixed Effects	Y,I		Y,I	
R^2	0.631		0.627	
N	3,749		3,749	

Table 21 presents regression results of the Cash Flow Prediction Model (Model 1) and the Certain and Uncertain Accrual Model (Model 2) based on the model of Barth et al. (2016) (Robustness Check II). Dependent variable is CFO in t+1. CFO is operating cash flow in t. ΔSFP^A is change in total receivables plus deferred tax assets minus the sum of accounts payable, accrued expenses, pension liability, income taxes payable, and deferred tax liability. ΔSFP^B is change in the sum of inventories, prepaid expenses, income tax refund, property, plant, and equipment, intangible assets, deferred charges, investments and advances-equity, and long-term pension assets minus deferred revenues. ΔSFP^A_UNCERTAIN (ΔSFP^A_CERTAIN) is the sum of those accruals contained in SFP^A that are (not) flagged as CAPs. ΔSFP^B_UNCERTAIN (ΔSFP^B_CERTAIN) is the sum of those accruals contained in SFP^B that are (not) flagged as CAPs. OACC is total accruals minus the sum of ΔSFP^A and ΔSFP^B. See Appendix A for detailed variable descriptions and Appendix D for further explanation. Each model is estimated with year- (Y) and industry- (I) fixed effects. ***, ** and * denotes significance at 1 %, 5 % and 10 % level, respectively.

Table 22
Robustness Check II – Importance and Specificity

Panel A: Estimation Window 10 Years

		SPECIFICITY (median split)		
		Low	High	Diff
IMPORTANCE (median split)	Low	0.066	0.070	0.004 *
		n = 807	n = 906	
	High	0.074	0.071	-0.002
		n = 1,029	n = 736	
	Diff	0.008 ***	0.001	

Panel B: Estimation Window 5 Years

		SPECIFICITY (median split)		
		Low	High	Diff
IMPORTANCE (median split)	Low	0.068	0.072	0.004 *
		n = 807	n = 906	
	High	0.077	0.075	-0.002
		n = 1,029	n = 736	
	Diff	0.009 ***	0.003	

Table 22 presents results from two-sided t-tests for differences in means of the mean absolute prediction error (MAPE). Panel A (Panel B) presents results by using the previous ten (five) years of data to obtain the coefficients to determine predicted cash flows using the cash flow prediction model of Barth et al. (2016). The sample is splitted by the median of IMPORTANCE and SPECIFICITY into 'low'- (below median) and 'high'-affected (above median) groups, respectively. MAPE is calculated as the absolute value of the difference of current operating cash flow (CFO) and predicted operating cash flow (PRED_CFO). Presented are the means of these four groups. See Appendix A for detailed variable descriptions. ***, ** and * denotes significance at 1 %, 5 % and 10 % level, respectively.

Third, to verify whether our results for H2 are driven by the slightly different empirical approach that we used compared to H1, we disaggregate further ACC_UNCERTAIN of Equation 5 into ACC_UNCERTAIN_IMP and ACC_UNCERTAIN_NOIMP. Whereas the former refers to the fitted values of uncertain accrual components that are important for a given firm (above the median of IMPORTANCE), the latter captures the fitted value of uncertain accruals that are less important (below the median of IMPORTANCE). The results are presented in Table 23. We find that the coefficient of ACC_UNCERTAIN_IMP is even lower than ACC_UNCERTAIN_NOIMP and that the difference is highly significant. Compared to that, although the coefficient of ACC_UNCERTAIN_NOIMP is still smaller than that of ACC_CERTAIN, this difference remains insignificant, indicating no significant difference in their persistence. As a result, accruals that are flagged as CAPs and are more important for a given firm, being the least persistent

components of accruals with respect to the predictability of future cash flows. Moreover, we find inconclusive results for SPECIFICITY (untabulated results).

Table 23
Robustness Check III – Persistence of Important Accruals

	Certain and Uncertain Accrual Model	
CFO	0.740	***
	(71.41)	
ACC_CERTAIN	0.146	***
	(5.81)	
ACC_UNCERTAIN_NOIMP	0.114	***
	(2.71)	
ACC_UNCERTAIN_IMP	0.062	***
	(2.80)	
ACC_OTHER	0.087	**
	(4.55)	
Constant	0.043	***
	(10.32)	
Fixed Effects	Y,I	
R^2	0.632	
N	4,260	

Table 23 presents regression results by disaggregating ACC_UNCERTAIN further into ACC_UNCERTAIN_NOIMP and ACC_UNCERTAIN_IMP (Robustness Check III). Dependent variable is CFO in t+1. CFO is operating cash flow in t. ACC_UNCERTAIN_NOIMP are the fitted values from those accruals that are flagged as CAPs and are less important (based on IMPORTANCE). ACC_UNCERTAIN_IMP are the fitted values from those accruals that are flagged as CAPs and are more important (based on IMPORTANCE). ACC_CERTAIN are the fitted values from those accruals that are not flagged as CAPs and the intercept. ACC_OTHER are the fitted values from all other accruals and the residuals from Equation 4. See Appendix A for detailed variable descriptions. The model is estimated with year- (Y) and industry- (I) fixed effects. ***, ** and * denotes significance at 1 %, 5 % and 10 % level, respectively.

In our fourth set of robustness checks, we use the previous five years of data to obtain the coefficients for determining predicted cash flows instead of using the previous ten years. Accordingly, we verify whether our results are driven by the period for our estimation window of the in-sample regressions. The results are presented in Panel A of Table 24. As shown, all major inferences regarding our second hypothesis remain qualitatively unchanged. Moreover, we conduct a sensitivity test regarding our IMPORTANCE measure. To do this, we re-estimate IMPORTANCE as the sum of all uncertain accrual components

divided by total assets. The results of our 2x2 matrix are presented in Panel B of Table 24. As can be seen, our alternative IMPORTANCE measure does not change our results.

Table 24
Robustness Check IV – 5 Year Estimation Window and Alternative Measurement of Importance

Panel A: Estimation Window 5 Years

		SPECIFICITY (median split)		
		Low	High	Diff
IMPORTANCE (median split)	Low	0.030	0.035	0.005 ***
		n = 763	n = 1,367	
	High	0.042	0.044	0.002
		n = 1,362	n = 768	
	Diff	0.012 ***	0.009 ***	

Panel B: Alternative Measure of Importance

		SPECIFICITY (median split)		
		Low	High	Diff
IMPORTANCE (median split)	Low	0.030	0.072	0.004 *
		n = 763	n = 1,367	
	High	0.042	0.075	-0.002
		n = 1,362	n = 768	
	Diff	0.009 ***	0.003	

Table 24 presents results from two-sided t-tests for differences in means of the mean absolute prediction error (MAPE) of robustness check IV. The sample (n = 4,260) is splitted by the median of IMPORTANCE and SPECIFICITY into 'low'- (below median) and 'high'-affected (above median) groups, respectively. Panel A presents results from two-sided t-tests for differences in means of MAPE by using an estimation window of five years instead of ten years to estimate coefficients from the pooled regressions estimated each year between 2000 and 2014. Panel B presents results from two-sided t-tests for differences in means of MAPE by calculating IMPORTANCE as the sum of uncertain accruals divided by total assets. MAPE is calculated as the absolute value of the difference of current operating cash flow (CFO) and predicted operating cash flow (PRED_CFO). Presented are the means of these four groups. Two tailed t-statistics are presented in parentheses. See Appendix A for detailed variable descriptions. ***, ** and * denotes significance at 1 %, 5 % and 10 % level, respectively.

Fifth, our results may be driven by aggregating the fitted values of all accruals that are (not) flagged as CAPs. Hence, in our fifth robustness check (V), instead of using the fitted values, we classify each single accrual component of Equation 4 either as certain or uncertain (i.e., ΔAR_CERTAIN and ΔAR_UNCERTAIN). Afterwards, we include both forms of each accrual component in our main model. This procedure allows us to directly analyze the effect of measurement uncertainties within single accruals on their persistence with respect to future cash flows. We present results in Table 25. As can be seen, all uncertain accrual components, except ΔINV_UNCERTAIN, are less persistent than their

counterparts. We conclude that the predictive power of each accrual component is significantly lower if it is flagged as uncertain.

Table 25
Robustness Check V – Persistence of Single Uncertain Accruals

	Certain and Uncertain Accrual Model	
CFO	0.758	***
	(70.70)	
ΔAR_CERTAIN	0.408	***
	(12.71)	
ΔAR_UNCERTAIN	0.268	***
	(6.94)	
ΔINV_CERTAIN	0.217	***
	(4.07)	
ΔINV_UNCERTAIN	0.266	***
	(6.60)	
ΔAP_CERTAIN	-0.531	***
	(-12.23)	
ΔAP_UNCERTAIN	-0.516	***
	(-5.80)	
DEPR_CERTAIN	0.120	***
	(2.74)	
DEPR_UNCERTAIN	0.070	**
	(2.01)	
OTHER	0.198	***
	(12.96)	
Constant	0.035	***
	(8.45)	
Fixed Effects	Y,I	
R^2	0.648	
N	4,260	

Table 25 presents regression results by classifying each accrual component as either 'certain' or 'uncertain' (Robustness Check V). Dependent variable is CFO in t+1. CFO is operating cash flow in t. ΔAR_UNCERTAIN (ΔAR_CERTAIN) is the year-to-year change in accounts receivables if accounts receivables are (not) classified as CAPs. ΔINV_UNCERTAIN (ΔINV_CERTAIN) is the year-to-year change in inventories if inventories are (not) classified as CAPs. ΔAP_UNCERTAIN (ΔAP_CERTAIN) is the year-to-year change in accounts payable if accounts payables are (not) classified as CAPs. DEPR_UNCERTAIN (DEPR_CERTAIN) is depreciation, amortization and depletion if a firm classified its depreciation, amortization or depletion (not) as CAPs. OTHER is the difference between total accruals and ΔAR, ΔINV, ΔAP and DEPR. The model is estimated with year- (Y) and industry- (I) fixed effects. See Appendix A for detailed variable descriptions. ***, ** and * denotes significance at 1 %, 5 % and 10 % level, respectively.

Sixth, in our main analysis, we only consider firms with complete time series data. Thus, our results might be driven by 'survivorship bias'. Therefore, to verify whether our results are in fact driven by this limitation, we include all observations with sufficient financial data, omit the restriction of complete time series data, and re-perform our entire analysis. This increases our final sample to 5,456 firm-year observations. As can be seen in Table 26 and Table 27, our main results remain qualitatively unchanged.

Table 26

Robustness Check VI – Sample Selection without the Restriction of Complete Time Series Data – Accrual Persistence

	Model (1)		Model (2)	
	Cash Flow Prediction Model		Certain and Uncertain Accrual Model	
CFO	0.749	***	0.732	***
	(77.12)		(78.92)	
ΔAR	0.246	***		
	(9.62)			
ΔINV	0.214	***		
	(6.85)			
ΔAP	-0.454	***		
	(-12.20)			
DEPR	0.105	**		
	(3.38)			
OTHER	0.159	***		
	(11.84)			
ACC_CERTAIN			0.075	***
			(3.33)	
ACC_UNCERTAIN			0.035	*
			(1.73)	
ACC_OTHER			0.059	***
			(3.55)	
Constant	0.034	***	0.041	**
	(9.18)		(11.02)	
Fixed Effects	Y,I		Y,I	
R^2	0.631		0.615	
N	5,456		5,456	

Table 26 presents regression results of the Cash Flow Prediction Model (Equation 3) (Model 1) and the Certain and Uncertain Accrual Model (Equation 5) (Model 2) using a sample without the restriction of complete time series data (n = 5,456). Dependent variable is CFO in t+1. CFO is operating cash flow in t. ΔAR is the year-to-year change in accounts receivables. ΔINV is the year-to-year change in inventories. ΔAP is the year-to-year change in accounts payable. DEPR is depreciation, amortization and depletion. OTHER is the difference of total accruals and ΔAR, ΔINV, ΔAP and DEPR. ACC is earnings minus operating cash flows. ACC_UNCERTAIN are the fitted values from those accruals that are flagged as CAPs. ACC_CERTAIN are the fitted values from those accruals that are not flagged as CAPs and the intercept. ACC_OTHER are the fitted values from all other accruals and the residuals from Equation 4. Each model is estimated with year- (Y) and industry- (I) fixed effects. See Appendix A for detailed variable descriptions. ***, ** and * denotes significance at 1 %, 5 % and 10 % level, respectively.

Table 27
Robustness Check VI – Sample Selection without the Restriction of Complete Time Series Data – Importance and Specificity

Panel A: Mean Absolute Prediction Error (MAPE)

		SPECIFICITY (median split)		
		Low	High	Diff
IMPORTANCE (median split)	Low	0.033	0.038	0.005 ***
		n = 935	n = 1,793	
	High	0.042	0.043	0.001
		n = 1,769	n = 959	
	Diff	0.009 ***	0.005 ***	

Panel B: Mean Absolute Error Term (MAET)

		SPECIFICITY (median split)		
		Low	High	Diff
IMPORTANCE (median split)	Low	0.028	0.031	0.004 *
		n = 935	n = 1,793	
	High	0.034	0.033	-0.001
		n = 1,769	n = 959	
	Diff	0.006 ***	0.002 **	

Table 27 presents results from two-sided t-tests for differences in means of the mean absolute prediction error (MAPE) (Panel A) and mean absolute error term (MAET) (Panel B) using a sample without the restriction of complete time series data. The sample (n = 5,456) is splitted by the median of IMPORTANCE and SPECIFICITY into 'low'- (below median) and 'high'-affected (above median) groups, respectively. MAPE is calculated as the absolute value of the difference of current operating cash flow (CFO) and predicted operating cash flow (PRED_CFO). MAET is calculated as the mean absolute error term estimated by the absolute value of the residuals of the cash flow prediction model (Equation 3) Presented are the means of these four groups. See Appendix A for detailed variable descriptions. ***, ** and * denotes significance at 1 %, 5 % and 10 % level, respectively.

8 Limitations and Future Research Suggestions

Our research should be considered in the light of some limitations. Our sample is based on the S&P 500 composition, which includes the largest corporations in the U.S. Therefore, our results might not be applicable to smaller stock corporations, and our findings might be driven by focussing only on larger firms. Consequently, future research might analyze the usefulness of CAP disclosures provided by smaller corporations with respect to measurement uncertainties embedded in financial statements. Having said that, there are huge discrepancies in the number and types of CAPs. In our main analysis, we focus only on those that relate to four distinct accrual components. Nevertheless, there are additional accruals that are also related to future cash flows and might be classified as a CAP, but are not considered in our main analysis. While we address this in an additional

analysis, future research may develop another empirical approach in order to consider the entire number of CAPs provided by a firm. Moreover, the models of Barth et al. (2001) and Barth et al. (2016) only consider accruals that align to prior, current, and next-period cash flows. Most firms have long-term accruals. Therefore, the underlying uncertainty resulting from the measurement process may not only affect a firm's next period cash flow, but also cash flows in multiple (future) periods. While we assess this in our first robustness check, future research could distinguish explicitly between CAPs that relate to short-term and long-term accruals to assess their predictive value with operating cash flows. By doing so, it would be possible to analyze how the difference between long-term and short-term accruals classified as CAPs affect our main inferences. Furthermore, we focus on whether CAPs contain *any* information about measurement uncertainties. However, to this day, it remains unknown whether CAP disclosures provide *new* information to analysts, investors, and other financial statement users. As stated before, CAP disclosures reflect proprietary information about measurement uncertainties within financial statements that is not available through other public channels. This may increase the relevance of CAP disclosures, because it provides detailed information about the measurement process of highly uncertain financial statement items and their consequences for a firm's financial performance. Therefore, future research could analyze whether CAP disclosures are useful for improving cash flow and earnings forecasts.

9 Conclusion

We shed light on the role of CAP disclosures to provide valuable information about firm-specific measurement uncertainties within accruals, as well as the overall reliability of accounting estimates. The primary research question in this study addresses whether and how CAP disclosures provide information about the persistence of specific accruals with respect to future cash flows. We provide initial evidence that those accruals flagged as CAPs are less useful in predicting future cash flows. Thus, CAP disclosures might be informative with respect to the subjectivity and uncertainties within accruals. Based on our empirical approach, we further find that the predictive power of uncertain accrual components (those that are flagged as CAPs) is not lower *per se*; it also depends mainly on their importance and to some extent on the specificity for a given firm. Our findings underline the importance of corporate disclosures for communicating measurement uncertainties in financial reports. Prior studies find that accruals based on a higher degree

of estimation are less persistent with respect to future earnings (e.g., Chen et al. 2019). Hence, it is necessary that investors, analysts, and other financial statement users are able to differentiate between accruals that are uncertain and those that are not susceptible to estimation errors. Our results show that CAP disclosures indeed convey such information.

Appendix A: Variable Descriptions

Variable	Description
ACC	Total accruals proxied as the difference between earnings and operating cash flow scaled by total assets (Source: Worldscope).
ACC_CERTAIN	The fitted value from those accruals that are flagged as CAPs.
ACC_UNCERTAIN	The fitted value from those accruals that are not flagged as CAPs.
ACC_UNCERTAIN_IMP	The fitted value from those accruals that are flagged as CAPs and are less importance (below median of IMPORTANCE).
ACC_UNCERTAIN_NOIMP	The fitted value from those accruals that are flagged as CAPs and are more importance (above median of IMPORTANCE).
ACC_OTHER	The fitted value from all other accruals and the residuals from Equation 4.
CAP_ACCRUAL	The number of accrual components of the Cash Flow Prediction Model of Barth et al. (2001) that are classified as CAPs.
CAP_OTHER	The number of all other CAPs proxied as the difference of CAP_TOTAL and CAP_ACCRUAL.
CAP_TOTAL	The total number of CAPs included in a firm's MD&A (Source: hand-collection).
CFO	Operating cash flow in t scaled by total assets (Source: Worldscope).
CFO3Y	The sum of CFO from t+1 to t+3.
DEPR	The sum of depreciation, amortization and depletion scaled by total assets (Source: Worldscope).
DEPR_CERTAIN	Depreciation, amortization and depletion scaled by total assets if depreciation, amortization and depletion are not classified as CAPs.
DEPR_UNCERTAIN	Depreciation, amortization and depletion scaled by total assets if depreciation, amortization and depletion are classified as CAPs.
IMPORTANCE	Importance of uncertain accrual components proxied as the sum of all uncertain accrual components scaled by the sum of all accrual components using the cash flow prediction model of Barth et al. (2001) (Source: Worldscope).
IND	Indicator variables for industry-fixed effects based on the Fama/French 12 industry portfolio (Source: Worldscope).
MAET	Mean absolute error term as the absolute value of the residuals from the cash flow prediction model of Barth et al. (2001).
MAPE	Mean absolute prediction error as the difference between the predicted cash flow and actual cash flow (PRED_ERR).
OACC	ACC minus the sum of ΔSFP^{A} and ΔSFP^{B} (Source: Worldscope).
OTHER	The difference between total accruals (ACC) and disaggregated accrual components (ΔAR, ΔINV, ΔAP, ΔINV, and DEPR) (Source: Worldscope).
PRED_CFO	Predicted cash flow in t+1 as proxied by multiplying the coefficients from the cross-sectional cash flow prediction model in t-1 with the related accounting numbers in t.
PRED_ERR	The difference between the actual and predicted cash flow.

Variable	Description
SPECIFICITY	The mean of the specificity score of each uncertain accrual component (SPECIFICITY_AR, SPECIFICITY_INV, SPECIFICITY_AP, SPECIFICITY_DEPR) based on a common term-weighting scheme from the textual analysis literature, namely term frequency-inverse document frequency (*tf-idf*) (Lougrahn and McDonald 2011). $$SPECIFICITY = \begin{cases} \frac{(1+log(cap\_i))}{(1+log(i))} \times log\frac{N}{cap\_n} & if\ cap\_i \geq 1 \\ 0 & otherwise \end{cases}$$ *with:* cap_i — raw count of the number of firms in a given year and industry that flagged AR, INV, AP, or DEPR as a CAP. cap_n — number of firms in a given year that flagged their AR, INV, AP, or DEPR as a CAP. N — total number of firms in a given year. i — total number of firms in a given year and industry.
YEAR	Indicator variables for year-fixed effects.
ΔAP	The year-to-year change in accounts payable scaled by total assets (Source: Worldscope).
ΔAP_CERTAIN	The year-to-year change in accounts payables scaled by total assets if accounts payables are not classified as CAPs.
ΔAP_UNCERTAIN	The year-to-year change in accounts payables scaled by total assets if accounts payables are classified as CAPs.
ΔAR	The year-to-year change in accounts receivables scaled by total assets (Source: Worldscope).
ΔAR_CERTAIN	The year-to-year change in accounts receivables scaled by total assets if accounts receivables are not classified as CAPs.
ΔAR_UNCERTAIN	The year-to-year change in accounts receivables scaled by total assets if accounts receivables are classified as CAPs.
ΔINV	The year-to-year change in inventories scaled by total assets (Source: Worldscope).
ΔINV_CERTAIN	The year-to-year change in inventories scaled by total assets if inventories are not classified as CAPs.
ΔINV_UNCERTAIN	The year-to-year change in inventories scaled by total assets if inventories are classified as CAPs.
ΔSFP^{A}	Total receivables plus deferred tax assets minus the sum of accounts payable, accrued expenses, pension liability, income taxes payable, and deferred tax liability (Source: Worldscope).
ΔSFP^{A}_CERTAIN	The sum of accruals contained in SFP^{A} that are not classified as CAPs.
ΔSFP^{A}_UNCERTAIN	The sum of accruals contained in SFP^{A} that are classified as CAPs.
ΔSFP^{B}	The sum of inventories, prepaid expenses, income tax refunds, property, plant and equipment, intangible assets, deferred charges, investments and advances-equity, and long-term pension assets minus deferred revenue (Source: Worldscope).
ΔSFP^{B}_CERTAIN	The sum of accruals contained in SFP^{B} that are not classified as CAPs.
ΔSFP^{B}_UNCERTAIN	The sum of accruals contained in SFP^{B} that are classified as CAPs.

Appendix B: Additional Explanations Empirical Framework Analysis 1

The following example will clarify our procedures to model 'certain' and 'uncertain' accruals:

1) First, we verify which accruals components are flagged as CAPs. For instance, a firm classifies its accounts payables (ΔAP) and depreciation, amortization and depletion (DEPR) as CAPs. Therefore, we argue that both accrual components are affected by measurement uncertainties and thus, are 'uncertain'. Consequently, we assume that a firm's accounts receivables (ΔAR) and inventories (ΔINV) are unaffected by measurement uncertainties, and thus are 'certain'.

2) Second, we run Equation 4. Afterwards, we use the estimated coefficients and aggregate the fitted values of the contemporaneous accounts receivables and inventory variables that capture accruals without measurement uncertainties as ACC_CERTAIN and the fitted values of the accounts payables and depreciation, amortization, and depletion variables that capture accruals with measurement uncertainties as ACC_UNCERTAIN. Moreover, we allocate OTHER as well as the error term into ACC_OTHER because we cannot assign single CAPs to these variables. The model takes the following form:

Example:

$$ACC_{i,t} = \underbrace{\alpha_0 + \beta_1 \Delta AR_{i,t} + \beta_2 \Delta INV_{i,t}}_{\text{ACC\_CERTAIN}} + \underbrace{\beta_3 \Delta AP_{i,t} + \beta_4 DEPR_{i,t}}_{\text{ACC\_UNCERTAIN}} + \underbrace{\beta_6 OTHER_{i,t} + \varepsilon}_{\text{ACC\_OTHER}}$$

3) We re-perform this procedure for each observation of our sample.

Appendix C: Additional Explanations Empirical Framework Analysis 2

The following general example illustrates our prediction procedures (example to forecast 2002 operating cash flow).

1) First, we estimate cross-sectionally the following regression using the previous ten years of data (spanning the period 1991 – 2000).

$$CFO_{i,t+1} = \alpha_0 + \beta_1 CFO_{i,t} + \beta_2 \Delta AR_{i,t} + \beta_3 \Delta INV_{i,t} + \beta_4 \Delta AP_{i,t} + \beta_5 DEPR_{i,t} + \beta_6 OTHER_{i,t} + \varepsilon \quad \text{(C1)}$$

2) Second, we obtain the estimated coefficients α_0 and β_i. Those coefficients are then used to predict firm-specific cash flows in 2002 by multiplying the independent variables in 2001 with the previously determined industry-specific estimated coefficients.

$$PRED\_CFO_{i,2002} = \alpha_0 + \beta_{1(2000)} CFO_{i,2001} + \beta_{2(2000)} \Delta AR_{i,2001} + \beta_{3(2000)} \Delta INV_{i,2001} + \beta_{4(2000)} \Delta AP_{i,2001} + \beta_{5(2000)} DEPR_{i,2001} + \beta_{6(2000)} OTHER_{i,2001} + \varepsilon \quad \text{(C2)}$$

3) Third, we determine the prediction error for each firm by comparing the estimated future cash flows with the actual cash flow. Therefore, we remain with predicted cash flow values from 2002 to 2014.

$$PRED\_ERR_{i,2002} = CFO_{i,2002} - PRED\_CFO_{i,2002} \quad \text{(C3)}$$

Appendix D: Cash Flow Prediction Model Barth et al. (2016)

The general model of Barth et al. (2016) takes the following form:

$$CFO_{i,t+1} = \alpha_0 + \beta_1 CFO_{i,t} + \beta_2 \Delta SFP^A_{i,t} + \beta_3 \Delta SFP^B_{i,t} + \beta_4 OACC_{i,t} + \beta_j \Sigma_j IND_i + \beta_k \Sigma_k YEAR_t + \varepsilon \quad \text{(D1)}$$

SFP^A is the sum of total receivables and deferred tax assets minus the sum of accounts payable, accrued expenses, pension liability, income taxes payable, and deferred tax liability. SFP^B is the sum of inventories, prepaid expenses, income tax refund, property, plant and equipment, intangible assets, deferred charges, investments and advances-equity, and long-term pension assets minus deferred revenue. ΔSFP^A (ΔSFP^B) is the change of SFP^A (SFP^B). OACC are those accruals other than SFP^A and SFP^B.

Following the procedure of our main analysis, we classify each accrual component that is included in SFP^A and SFP^B (see Appendix B) either as certain or uncertain using CAP disclosures:

1) In a first step, we regress total accruals (ACC) on the disaggregated accrual components of the Barth et al. (2016) model.

2) In a second step, we disaggregate both accrual components into 'certain' and 'uncertain'. SFP^A_UNCERTAIN and SFP^B_UNCERTAIN (SFP^A_CERTAIN and SFP^B_CERTAIN) are the sum of the fitted values of those accruals contained in SFP^A and SFP^B that are (not) classified as CAPs. The following example will clarify our procedure: For instance, a firm classifies its pension liabilities, accounts payables, inventories, intangible assets, and deferred revenue as CAPs. Thus, our main variables are estimated as the sum of the fitted values from the following variables:

SFP^A_CERTAIN	=	Accounts receivables + deferred tax assets – (accrued expenses + income taxes payables + deferred tax liability)
SFP^A_UNCERTAIN	=	Accounts payable + pension liability
SFP^B_CERTAIN	=	Prepaid expenses + income tax refund + property, plant, and equipment + deferred charges + investments and advances equity + long-term pension assets
SFP^B_UNCERTAIN	=	(Inventories + intangible assets) – deferred revenues

3) In a third step, we replace the disaggregated earning components of the Barth et al. (2016) model and run the following regression:

$$\begin{aligned} CFO_{i,t+1} = {} & \alpha_0 + \beta_1 CFO_{i,t} + \beta_2 \Delta SFP\_CERTAIN^{A}_{i,t} \\ & + \beta_3 \Delta SFP\_UNCERTAIN^{A}_{i,t} + \beta_4 \Delta SFP\_CERTAIN^{B}_{i,t} \\ & + \beta_5 \Delta SFP\_UNCERTAIN^{B}_{i,t} + \beta_6 OACC_{i,t} + \beta_j \Sigma_j IND_i \\ & + \beta_k \Sigma_k YEAR_t + \varepsilon \end{aligned} \quad \text{(D2)}$$

CHAPTER 3:
INSTITUTIONAL CORRELATES WITH FEMALE BOARD REPRESENTATION

Abstract

Gender equality on boards is a global and highly politicized issue. To this day, there is considerable cross-country variation in female board representation. We examine institutional supply- and demand-side factors associated with this issue. Our results indicate that a societal climate of gender equality contributes to more women on boards, mainly through fostering the supply of suitable candidates. Therefore, the glass ceiling should be improved through a societal supply-side effort which needs to complement demand-side (quota) regulation.

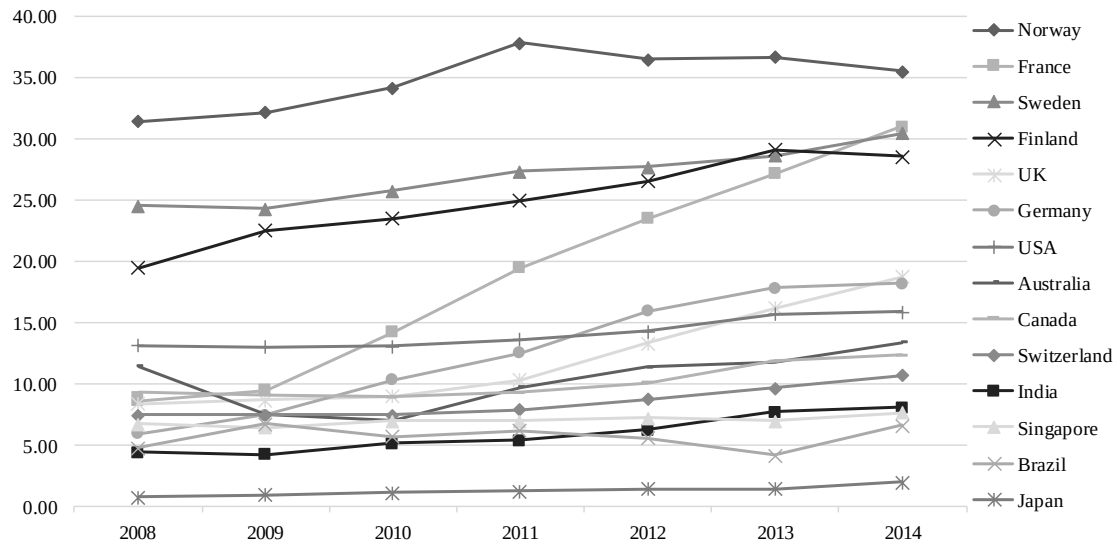
This part of the thesis is a joint project with Thomas R. Loy. Chapter 3 has been published as Loy, T., Rupertus, H. (2018), Institutional Correlates with Female Board Representation. *Finance Research Letters*, 24, 238-246. DOI: 10.1016/j.frl.2017.09.013.

Acknowledgements: We gratefully acknowledge financial support by FACT Alumni Universität Bayreuth e.V. We also thank Florian Federsel and Marie-Thérèse Meyer for their excellent research assistance.

1 Introduction

Female corporate board representation is an increasingly important and highly politicized issue. Several European countries obligate a gender quota on corporate boards to foster female participation in economic activity. In 2008, Norway adopted the first mandatory gender quota regulation. Similar requirements have been or will be adopted in Belgium, France, Germany, and the Netherlands over the next years (Terjesen et al. 2015). Further European countries will likely follow in response to EU initiatives (European Commission 2016). Internationally, gender quotas are stipulated in a wide range of voluntary corporate governance codes (Terjesen et al. 2015). However, despite these efforts, there still are considerable differences in terms of average female board membership (Figure 7).

Figure 7
Country-Level Average Percentage of Women on Boards of Directors



On the high end, Scandinavian firms exhibit female board representation of about 30 percent. Contrariwise, the ratio for firms from Brazil, India, Japan, or Singapore is less than 10 percent. So far, the academic literature has focused on the business case for board gender diversity (e.g., Adams and Ferreira 2009; for meta-analyses: Pletzer et al. 2015; Post and Byron 2015). Boards as corporations' upper echelons enhance firm performance, in large part, through interaction and sharing of knowledge and resources (Hambrick 2007). Proponents argue that gender diversity improves the quality of board discussions, attributable to more creativity (Nemeth 1986), and a wider range of perspectives (Hillman et al. 2007). Critics point to more conflicts between board members which arise

from internal divisions and social categorization tendencies (Gul et al. 2011; Tajfel 1979). As such, prior studies provide mixed results regarding the association of board gender diversity and firms' financial performance. Even more problematically, most studies focus on single countries (e.g., Adams and Ferreira 2009; Farrell and Hersch 2005; Carter et al. 2003; Huang and Kisgen 2013; Gul et al. 2011). However, to make a convinced statement about the effect of female board representation on e.g., firm performance, future research should control for aforementioned global differences in gender equality as well as the underlying reasons.

Yet, surprisingly little academic research tries to answer the question why in some countries women are far more underrepresented in the boardroom than in others. As such, we contribute to this ongoing debate and answer a call for additional cross-country research (Gabaldon et al. 2016) by systematically and empirically disentangling institutional supply- and demand-side factors associated with the issue. Our results exhibit that longer-term supply-side factors need to complement short-term demand-side regulation in order to crack the glass ceiling. Hence, our paper contributes to the ongoing societal and political discussion revolving around gender equality, corporate governance, and the glass ceiling.

2 Theoretical Foundations and Research Questions

Post and Byron (2015) present a meta-analysis of gender diversity and firm performance and find two mediating factors. First, they argue that increased shareholder protection and directors' legal liability result in improved consideration of female directors' experiences, knowledge, and values. If dissenting voices are wilfully dismissed but later proven to be correct, this will most likely result in repercussions against stereotyping directors. As such, it is in the best interest of their (male) colleagues to regard female cognitive frames as an advantage in sound decision-making. Second, there has to be a societal climate that enables women to acquire equal skills, education, and human capital to fulfil their fiduciary role on corporate boards. Otherwise female socio-economic disadvantages render the effects of greater diversity obsolete. Moreover, a greater societal gender equality results in companies requiring female directors to gain legitimacy (e.g., Bear et al. 2010).

Despite the general appeal of these theories, there is surprisingly little conversation in the literature why these considerable country-level differences exist in the first place. A rather

small line of research tries to explore some country-level factors that facilitate women entering boardrooms. It attributes most of the effect to fundamental legal and cultural institutions (Grosvold and Brammer 2011), emphasizes (smaller) gender pay gaps and the proportion of women in (middle) management, as a pool of potential board candidates (Terjesen and Singh 2008), focuses on actions by individual female politicians and their interplay with political parties, business associations, and other stakeholders (Seierstad et al. 2017), analyzes the determinants of (voluntary) gender quota regulation (Terjesen et al. 2015), or determines a larger proportion of women working full-time to be a crucial prerequisite of female board representation (Adams and Kirchmaier 2015). However, there is no research systematically examining which supply- or demand-side levers could be most effective in cracking the glass ceiling.

We build on the notion that director selection is the result of a market process balancing firm needs, board dynamics, director characteristics as well as environmental and legal aspects (Withers et al. 2012). Thus, the glass ceiling likely has supply- as well as demand-side explanations. Therefore, we answer a call for research by systematically and empirically disentangling these factors. Gabaldon et al. (2016) highlight that this literature "would benefit from a more cross-cultural perspective, analyzing whether the gender gap on boards is due to supply or demand factors and how this varies across cultures" (Gabaldon et al. 2016, p. 381). We state the following research question:

***RQ**: Which supply- and demand-side factors are associated with the rate of female board representation on the country-level?*

3 Data and Estimation Strategy

In order to test our research question, we consider institutional correlates with average female corporate board presence (Equation 7), based on a minimum of 10 firm-year observations per country-year to achieve country-level averages unbiased by a small number of outliers (Source: Thomson Reuters *ASSET4* (GQBOARD)). This restriction results in a sample of 37 countries[41] or 418 country-year observations for the period 2002 through 2015, respectively.

[41] These are Australia, Austria, Belgium, Brazil, Canada, Chile, China, the Czech Republic, Egypt, Finland, France, Germany, Greece, India, Indonesia, Ireland, Israel, Italy, Japan, Korea (Rep.), Malaysia,

$$GQBOARD_{i,t} = \alpha_0 + \beta_1 GGGI_{i,t} + \beta_2 INVPROT_{i,t} + \beta_3 QUOTA_{i,t} + \beta_4 GOVCODE_{i,t} + \beta_5 GDPCAP_{i,t} + \beta_j \Sigma_j LEGOR_{i,t} + \beta_k \Sigma_k CONTINENT_{i,t} + \beta_l \Sigma_l YEAR_{i,t} + \varepsilon \quad (7)$$

Initially, we follow Post and Byron (2015) and predict that societal gender equality as well as minority investor protection positively influence female board participation. We employ *World Economic Forum*'s (WEF) Global Gender Gap Index (GGGI). It combines a total of 14 variables covering topics from basic needs, such as health and survival, to political empowerment, economic participation, and opportunity. Furthermore, *World Bank*'s strength of minority investor protection index (INVPROT) combines three indices covering (1) the extent and frequency of related party transactions disclosure, (2) the ability to sue over related party transactions, and (3) the ease of those shareholder suits. Moreover, we add international mandatory quota regulations (QUOTA) as well as voluntary corporate governance code (GOVCODE) stipulations (Terjesen et al. 2015). We predict positive signs for investor protection, gender equality, as well as gender-related board regulation. The natural logarithm of GDP per capita (GDPCAP) proxies for overall economic development. Legal orientation (LEGOR) clusters countries in terms of legal families. Although concrete legal frameworks may differ quite substantially across countries, basic premises rooted in common legal traditions still prevail (La Porta et al. 1998). Continent-fixed effects (CONTINENT) control for cultural heritage and path dependence in gender issues (Grosvold and Brammer 2011; Terjesen et al. 2015), as well as colonial history which still has a profound impact on social life, economic development, and education (Klerman et al. 2011). Finally, we add year-fixed effects (YEAR).

GGGI as an aggregate score does not differentiate between supply- and demand-factors. We employ confirmatory Principal Component Analysis (PCA) to extract two components which reflect underlying commonalities among the 14 individual indicators. We classify the ratio of females over males in the labor force (LABFORCE), the ratio of female over male senior officials and managers (MANAGERS), and the ratio of female over male professional and technical workers (PROFESSIONALS) as *supply* factors. Whereas, wage equality for similar work (WAGEEQUALITY) is likely associated with

Mexico, the Netherlands, New Zealand, Norway, the Philippines, Poland, the Russian Federation, Singapore, South Africa, Spain, Sweden, Switzerland, Thailand, Turkey, the U.K., and the U.S. Increasing the threshold to 20 (30) firm-year observations does not materially influence our results but decreases our sample size considerably.

the *demand* for female board members, estimated earned income (EARNINCOME) might both be a *supply* (i.e., females are more likely to apply if wages are higher) as well as *demand* (i.e., employers might assign a wage premium to women) factor. Moreover, we assign all four indicators covering educational attainment to the *supply*-side. These include the ratios of female over male enrolment in primary (PRIMARY), secondary (SECONDARY) and tertiary (TERTIARY) education, as well as female over male literacy (LITERACY). If women achieve relatively higher educational levels, firms can choose their directors out of a larger candidate pool. Additionally, the female to male sex ratio at birth (SEXRATIO) and the ratio of healthy male to female lifespans (LIFEEXPECTANCY), likely affect *demand*. Since board positions regularly require some seniority and firms also benefit from longer director tenure, their increased healthy life expectancy should disproportionately benefit women. On the one hand, women in parliament (WOMENPARL) and in ministerial positions (WOMENMIN) are a result of female political empowerment. They serve as role models for other women who aspire positions of power, hence fostering the *supply*. On the other hand, we assign a country's ratio of female to male heads of state over the last 50 years (FEMHEADSTATE) to the *demand*-side. We assume that firms are more open-minded to female board appointments if the population votes women into the position of utmost power. Figure 8 summarizes WEF's order by topic as well as our predictions with respect to *demand*- and *supply*-side effects of GGGI's underlying variables. Bold font indicates significant and material factor loadings (>.3) in a PCA. There are no significant and material side loadings on the respective other component.

Figure 8
Disaggregation of World Economic Forum's Global Gender Gap Index (GGGI) into its Underlying Supply and Demand Factors

Topic \ Supply-Demand	Code	Description	Supply	Demand
Economic Participation and Opportunity	LABFORCE	Females over males in labor force	X	
	WAGEEQUALITY	Wage equality for similar work		**X**
	EARNINCOME	Estimated earned income	**X**	**X**
	MANAGERS	Female over male senior officials	**X**	
	PROFESSIONALS	Female over male professional and technical workers	**X**	
Educational Attainment	LITERACY	Female over male literacy	**X**	
	PRIMARY	Female over male primary/secondary/tertiary education enrolment	X	
	SECONDARY		**X**	
	TERTIARY		**X**	
Health and Survival	SEXRATIO	Female to male sex ratio at birth		X
	LIFEEXPECTANCY	Male to female healthy life expectancy		**X**
Political Empowerment	WOMENPARL	Ratio of women in parliament	X	
	WOMENMIN	Ratio of women in ministerial positions	X	
	FEMHEADSTATE	Ratio of female to male heads of state		**X**

Out of the 14 variables, six supply-side variables load significantly on the first component while four demand-side variables load on the second component. As such, MANAGERS, PROFESSIONALS, LITERACY, SECONDARY, and TERTIARY predominantly determine the SUPPLY component. Whereas, WAGEEQUALITY, LIFEEXPECTANCY, and FEMHEADSTATE load significantly on the second (DEMAND) component. EARNINCOME significantly loads on both components, also in line with our earlier discussion.[42]

[42] The eigenvalues of both components are >1.0 and the Kaiser-Meyer-Olkin (KMO) criterion is >.8, indicating both suitability and good fit. Both components are essentially unrelated with a component rotation value of .0699. Significant and material factor loadings (>.3) are .367 for MANAGERS, .432 for PROFESSIONALS, .308 for LITERACY, .375 for SECONDARY, .403 for TERTIARY, .520 for WAGEEQUALITY, .512 for LIFEEXPECTANCY, and .354 for FEMHEADSTATE. EARNINCOME

Recent research also calls for addressing unobserved heterogeneity and simultaneity in board governance (Wintoki et al. 2012). As the GGGI, and its components, might itself be influenced by socio-economic and long-standing cultural values, we employ dynamic panel estimation (SYS-GMM), which also alleviates concerns of autocorrelation (Blundell and Bond 1998). As such, we include lagged GQBOARD as an additional control. SYS-GMM simultaneously estimates a system of equations in first differences and levels, in which the level equation includes time-invariant controls, such as legal orientation. We focus on a one-step estimation, as it entails similar efficiency compared to the two-step version, but provides more reliable estimates and unbiased standard errors in finite samples with shorter time-series (Soto 2009).

4 Results

4.1 Descriptive Statistics and Univariate Results

Average female board representation is 10.85 percentage points with a considerable standard deviation (Table 28). The GGGI has an average value of about 0.71 (out of a theoretical maximum of 1.0). Minority investor protection exhibits an average score of 6.21 (out of 10). About 12 (7.40) percent of the observations are subject to 'soft' GOVCODE and 'hard' gender quota regulation, respectively. More than half of the observations are European, with Asia being a distant second (16.0 %). In terms of legal origin, the French code-law (37.1 %) and Anglo-Saxon case-law (32.8 %) traditions dominate, followed by Germanic (20.0 %) and Scandinavian (10.0 %) code-law traditions.

loads significantly on the SUPPLY (.323) as well as the DEMAND (.368) component. No predetermined supply-side variables significantly load on the DEMAND component, and vice versa.

Table 28
Summary Descriptive Statistics

Variables	Mean	25 %	Median	75 %	SD	N
GQBOARD	10.85	5.470	8.850	14.56	7.830	418
GGGI	0.710	0.671	0.703	0.751	0.058	418
SUPPLY	0.128	-0.381	0.645	1.226	1.818	295
DEMAND	0.000	-1.072	-0.064	0.996	1.319	295
GENDERCULTURE	4.635	4.410	4.710	5.020	0.434	372
INVPROT	6.207	5.000	6.000	7.300	1.642	418
GDPCAP	10.21	9.911	10.39	10.60	0.603	418
QUOTA	0.074				0.262	418
GOVCODE	0.124				0.330	418
EUROPE	0.507				0.501	418
AFRICA	0.074				0.262	418
ASIA	0.160				0.367	418
AMERICA	0.136				0.343	418
AUSPAC	0.122				0.328	418
LEGOR_FR	0.371				0.484	418
LEGOR_UK	0.328				0.469	418
LEGOR_GE	0.200				0.401	418
LEGOR_SC	0.100				0.301	418

Table 28 presents summary descriptive statistics for all country-year observations for the period 2002 through 2015. GQBOARD is the average ratio of female board members per country-year. GGGI is an index value of the Global Gender Gap Index (World Economic Forum). SUPPLY is the first component of the principal component analysis of the 14 single scores which constitute GGGI. DEMAND is the second component of the principal component analysis of the 14 single scores which constitute GGGI. GENDERCULTURE is the value for cultural gender egalitarianism. INVPROT is an index value of the minority investor protection index compiled by the World Bank. GDPCAP is the natural logarithm of GDP per capita. QUOTA is an indicator variable equal to 1 if the country's non state-owned firms are subject to voluntary gender regulation, 0 otherwise. GOVCODE is an indicator variable equal to 1 if the country's non state-owned firms are subject to a mandatory gender quota, 0 otherwise. EUROPE, AFRICA, ASIA, AMERICA, and AUSPAC are indicator variables equal to 1 if the country is located in the respective continent, 0 otherwise. LEGOR is an indicator variable equal to 1 if the country follows a certain legal tradition (French (FR), UK, Germanic (GE), Scandinavian (SC)), 0 otherwise. For binary variables only means and standard deviations are reported. See Appendix A for detailed variable descriptions.

Next, we present univariate results. Consistent with prior evidence, GGGI and INVPROT are significantly and positively associated with GQBOARD (Post and Byron 2015). So are 'hard' and 'soft' board gender regulations (Terjesen et al. 2015), as well as our newly created institutional supply- and demand-side prime components (Table 29).

Table 29
Pearson Correlations

	Variable	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)	(15)	(16)	(17)	(18)
(1)	GQBOARD	1.00																	
(2)	GGGI	**0.68**	1.00																
(3)	SUPPLY	**0.47**	**0.75**	1.00															
(4)	DEMAND	**0.48**	**0.56**	**0.13**	1.00														
(5)	INVPROT	**0.19**	**0.16**	**0.13**	**0.16**	1.00													
(6)	GENDERCULTURE	**0.19**	**0.46**	**0.47**	0.03	0.03	1.00												
(7)	GDPCAP	**0.31**	**0.48**	**0.45**	**0.11**	**0.26**	**0.60**	1.00											
(8)	QUOTA	**0.40**	**0.09**	**0.13**	**0.12**	**0.18**	0.03	0.07	1.00										
(9)	GOVCODE	**0.27**	**0.26**	**0.19**	**0.16**	-0.02	**0.24**	**0.24**	**-0.11**	1.00									
(10)	EUROPE	**0.24**	**0.43**	**0.35**	0.09	**-0.30**	**0.51**	**0.48**	-0.05	**0.26**	1.00								
(11)	AFRICA	**0.18**	**-0.14**	-0.07	0.02	**0.12**	**-0.26**	**-0.22**	**0.41**	**-0.11**	**-0.29**	1.00							
(12)	ASIA	**-0.31**	**-0.37**	**-0.52**	0.03	0.02	**-0.37**	**-0.3**	-0.05	**-0.16**	**-0.44**	**-0.12**	1.00						
(13)	AMERICA	**-0.16**	**-0.12**	0.07	**-0.31**	**0.09**	**0.21**	**-0.09**	**-0.11**	**-0.15**	**-0.4**	**-0.11**	**-0.17**	1.00					
(14)	AUSPAC	0.02	-0.01	0.00	**0.14**	**0.24**	**-0.32**	**-0.12**	-0.08	0.04	**-0.38**	**-0.11**	**-0.16**	**-0.15**	1.00				
(15)	LEGOR_FR	**-0.27**	**-0.32**	**-0.15**	**-0.39**	**-0.44**	-0.08	**-0.29**	**-0.12**	**-0.17**	**0.11**	-0.07	**-0.12**	**0.17**	**-0.17**	1.00			
(16)	LEGOR_UK	**0.10**	0.07	**0.10**	**0.30**	**0.71**	**0.10**	0.07	**0.15**	-0.02	**-0.42**	**0.25**	-0.04	0.08	**0.41**	**-0.54**	1.00		
(17)	LEGOR_GE	**-0.23**	**-0.14**	**-0.16**	**-0.18**	**-0.27**	**-0.11**	**0.08**	**-0.14**	**0.15**	**0.11**	**-0.14**	**0.30**	**-0.20**	**-0.19**	**-0.38**	**-0.35**	1.00	
(18)	LEGOR_SC	**0.59**	**0.59**	**0.29**	**0.41**	-0.04	**0.13**	**0.25**	**0.15**	**0.09**	**0.33**	**-0.09**	**-0.15**	**-0.13**	**-0.12**	**-0.26**	**-0.23**	**-0.17**	1.00

Table 29 presents Pearson correlations. GQBOARD is the average ratio of female board members per country-year. GGGI is an index value of the Global Gender Gap Index (World Economic Forum). SUPPLY is the first component of the principal component analysis of the 14 single scores which constitute GGGI. DEMAND is the second component of the principal component analysis of the 14 single scores which constitute GGGI. INVPROT is an index value of the minority investor protection index compiled by the World Bank. GENDERCULTURE is the value for cultural gender egalitarianism. GDPCAP is the natural logarithm of GDP per capita. QUOTA is an indicator variable equal to 1 if the country's non state-owned firms are subject to voluntary gender regulation, 0 otherwise. GOVCODE is an indicator variable equal to 1 if the country's non state-owned firms are subject to a mandatory gender quota, 0 otherwise. EUROPE, AFRICA, ASIA, AMERICA, and AUSPAC are indicator variables equal to 1 if the country is located in the respective continent, 0 otherwise. LEGOR is an indicator variable equal to 1 if the country follows a certain legal tradition (French (FR), UK, Germanic (GE), Scandinavian (SC)), 0 otherwise. Bold font indicates significance at the two-tailed 10 %-level or higher.

4.2 Multivariate Results

Our results reveal several interesting insights. We are able to confirm the mediating effects of a societal climate of gender equality (GGGI) and minority investor protection (INVPROT) on board composition. The explanatory power of the OLS results is rather high with R^2 values between 73 and 76 percent. Investor protection is highly insignificant in the OLS specifications but significant in the SYS-GMM dynamic panel estimation which simultaneously controls for autocorrelation and endogeneity (Table 30). As previously discussed, the SYS-GMM results are likely more robust than cross-sectional methods for small samples.

In terms of GGGI, we exhibit that overall gender equality also spills over onto corporate boards. A one standard deviation increase in the GGGI represents a (highly) significant increase in the average percentage of female directors between 0.65 (SYS-GMM) and 2.68 (OLS) percentage points on the country-level. This confirms the signaling theory of board diversity (Bear et al. 2010). If a society implicitly expects gender equality, firms seem to cater to these expectations for legitimacy. As we substitute SUPPLY and DEMAND for aggregate GGGI, we encounter somewhat surprising results. While the SUPPLY component is significant, even in the stricter SYS-GMM setup, the DEMAND component also has the expected positive sign, yet it is only significant on a five percent level in OLS. Moreover, we show that gender quota regulations, in the form of mandatory, 'hard' regulations (QUOTA), and voluntary corporate governance code stipulations (GOVCODE) seem to incrementally contribute to more women on corporate boards, with the exception of the strictest specification (Model 4 of Table 30). In summary, this seems to suggest that highly regulated demand-side oriented legislation ought to be complemented by a societal focus on supply-side measures, such as education and giving women more opportunities to move into managerial and professional roles, in order to achieve their intended goals. More specifically, career progressions are oftentimes impaired by the stress to balance work and family life. These measures might, therefore, include increased opportunities to take advantage of childcare and improvements in paid maternity leave for mothers and fathers, alike (for a discussion of these issues, e.g., Grosvold and Brammer 2011; Adams and Kirchmaier 2015).

Table 30
Main Results

	Model 1	Model 2	Model 3	Model 4
	OLS	SYS-GMM	OLS	SYS-GMM
$GQBOARD_{t-1}$		0.808 ***		0.797 ***
		(10.23)		(8.32)
GGGI	44.63 ***	10.81 **		
	(5.11)	(2.42)		
SUPPLY			0.707 ***	0.288 **
			(2.71)	(2.31)
DEMAND			0.964 **	0.304
			(2.14)	(1.14)
INVPROT	0.234	0.249 ***	0.307	0.441 ***
	(0.98)	(2.66)	(0.91)	(3.23)
QUOTA	6.107 ***	2.024 **	5.524 ***	1.568
	(4.37)	(2.53)	(3.29)	(1.52)
GOVCODE	2.514 **	0.898 **	1.089	0.589
	(2.33)	(2.17)	(0.95)	(1.44)
GDPCAP	0.097	-0.006	-0.048	0.064
	(0.14)	(-0.02)	(-0.05)	(0.17)
LEGOR_FR	-9.600 ***	-1.541	-12.45 ***	-1.391
	(-4.59)	(-1.50)	(-7.32)	(-0.98)
LEGOR_UK	-9.080 ***	-2.059 **	-11.60 ***	-2.747 **
	(-4.06)	(-2.33)	(-5.58)	(-2.09)
LEGOR_GE	-10.65 ***	-1.687 *	-14.38 ***	-1.211
	(-5.40)	(-1.93)	(-8.17)	(-1.25)
EUROPE	-0.523	0.798	1.852	1.211
	(-0.34)	(1.58)	(0.81)	(1.81)
AFRICA	3.803 **	-0.191	2.474	-0.174
	(2.63)	(-0.29)	(1.04)	(-0.19)
ASIA	-2.797 *	-0.409	-1.601	-0.349
	(-1.85)	(-0.87)	(-0.71)	(-0.49)
AMERICA	-1.800	-0.239	-1.095	-0.124
	(-1.18)	(-0.62)	(-0.55)	(-0.23)
Constant	-19.38 **	omitted	17.89	omitted
	(-2.16)		(1.65)	
Fixed Effects	Y	Y	Y	Y
R^2	0.732	n/a	0.760	n/a
N	418	381	295	306
Arr.-Bond AR(1)		0.004		0.031
Arr.-Bond AR(2)		0.559		0.739
Hansen (p-value)		1.000		1.000

Table 30 presents multivariate results (Equation 7) for the OLS regressions (Model 1 and Model 3) and dynamic panel estimation models (SYS-GMM) (Model 2 and Model 4). See Appendix A for detailed variable descriptions. Standard errors are clustered at the country-level (two-tailed t-statistics (z-statistics) in parentheses). ***, **, and * denote significance at the 1 %, 5 %, and 10 % level, respectively.

Arrelano-Bond tests reveal significant autocorrelations with respect to one-period lags but not two-period lags. Given that supervisory board members serve multiple year terms it does not seem surprising that contemporaneous female board representation is largely determined by past realizations. A potential downside of the SYS-GMM models is that they might be weakened through a comparatively large number of instruments in relation to country-year observations. Nevertheless, Hansen tests for overidentification determine that the results are robust.

4.3 Additional Results

Much research is devoted on the association of culture and gender roles and gender stereotypes (e.g., Inglehart and Norris 2003). As (national) culture is comprised of long-standing traditions and, therefore, is highly path-dependent, our main results may reflect differences in culture rather than the effects of political initiatives geared towards empowering women in corporate life. To control for the impact of culture, we employ the country-level score for gender egalitarianism values (GENDERCULTURE) from the Globe project[43] (House 2004).

Prior research, building on sex segregation theory (Cejka and Eagly 1999; Glick 1991), suggests that different industries require certain levels of masculinity and, thus, are more likely to appoint women to leadership positions (e.g., Cumming et al. 2015). Therefore, we re-estimate our institutional analyses on the firm-level and add industry-fixed effects (building on the Fama-French 12 industry-framework) to Equation 7 (Table 31). These analyses also employ lagged values of GQBOARD as an additional control, but build on OLS estimation, since SYS-GMM is less efficient and likely biases estimates for larger samples.[44] Both additional analyses confirm our main results. GGGI as well as SUPPLY remain (highly) significant and are virtually unchanged. Our results are robust to the addition of cultural gender egalitarianism (GENDERCULTURE) or industry-fixed effects.

[43] http://globeproject.com/

[44] In contrast to the country-level analysis, with at most 418 country-year observations, the additional firm-year analyses have sample sizes of 25,646 (21,205) firm-year observations, respectively.

Table 31
Additional Results

	Country-Level		Firm-Level	
Method	Model 1	Model 2	Model 3	Model 4
$GQBOARD_{t-1}$	0.818 ***	0.809 ***	0.870 ***	0.873 ***
	(8.98)	(7.53)	(184.71)	(179.12)
GGGI	10.19 **		6.359 ***	
	(2.16)		(4.00)	
SUPPLY		0.264 **		0.176 ***
		(2.21)		(4.71)
DEMAND		0.319		-0.059
		(1.09)		(-0.97)
GENDERCULTURE	0.252	-0.123		
	(0.43)	(-0.17)		
INVPROT	0.283 **	0.475 ***	0.061	0.112 **
	(2.43)	(2.59)	(1.51)	(1.97)
QUOTA	2.461 ***	2.247 **	1.082 ***	0.898 **
	(3.24)	(2.04)	(4.92)	(2.51)
GOVCODE	0.896 *	0.652	1.087 ***	0.804 ***
	(1.72)	(1.51)	(8.23)	(5.71)
GDPCAP	-0.056	0.127	0.356 ***	0.885 ***
	(-0.21)	(0.27)	(3.70)	(5.01)
LEGOR_FR	-1.568	-1.183	-0.909 ***	-0.822 ***
	(-1.49)	(-0.73)	(-3.06)	(-2.69)
LEGOR_UK	-2.305 **	-2.733 *	-1.318 ***	-1.564 ***
	(-2.34)	(-1.77)	(-5.28)	(-5.61)
LEGOR_GE	-1.759 *	-1.695	-1.970 ***	-2.188 ***
	(-1.91)	(-1.16)	(-6.94)	(-7.10)
EUROPE	0.555	1.208	0.657 ***	0.570 ***
	(0.89)	(1.59)	(4.41)	(3.01)
AFRICA	-0.473	-0.548	0.664 **	1.128 ***
	(-0.64)	(-0.50)	(2.54)	(3.33)
ASIA	-0.476	-0.272	-0.446 *	-0.813 ***
	(-0.88)	(-0.33)	(-2.14)	(-2.89)
AMERICA	-0.337	0.087	0.298 **	-0.317 *
	(-0.64)	(0.13)	(2.00)	(-1.74)
Constant	omitted	omitted	-5.443 ***	-6.317 ***
			(-3.68)	(-3.34)
Fixed Effects	Y	Y	Y, I	Y, I
R^2	n/a	n/a	0.816	0.821
N	339	246	25,646	21,205
Arr.-Bond AR(1)	0.012	0.046		
Arr.-Bond AR(2)	0.709	0.424		
Hansen (p-value)	1.000	1.000		

Table 31 presents additional multivariate results (Equation 7) for the dynamic panel estimation model (SYS-GMM) (Model 1 and 2) and OLS regression (Model 3 and Model 4). Standard errors are clustered at the country- (firm-) level (two-tailed t-statistics (z-statistics) in parentheses). ***, **, and * denote significance at the 1 %, 5 %, and 10 % level, respectively.

5 Conclusion

Over the last decades, women have made significant advances in higher education, political activism, as well as labor force participation. Despite these efforts, in most countries men still dominate top-level corporate positions by large margins. With respect to country-level, institutional correlates, we show that functioning outside investor protection as well as a societal climate of gender equality contribute to increased female board participation. Additionally, we provide initial evidence that supply-side factors seem to be necessary to complement mandatory as well as voluntary gender quota regulation.

Going forward, targeting board compositions of listed corporations through regulation may not be sufficient in itself, if there is insufficient supply of qualified women in the workforce. Therefore, our paper presents a societal case for gender fairness which goes above and beyond (sometimes) rather symbolic short-term fixes. While gender quotas might regulate the *demand* for female upper echelons, they likely are ineffective to increase the *supply* of suitable candidates in the short run. Legislators should rather focus on supply-side measures, such as education and giving women more opportunities to move into managerial and professional roles.

Appendix A: Variable Descriptions

Variable	Description
CONTINENT	Indicator variable equal to 1 if the country is located on the respective continent (e.g., EUROPE, AFRICA, ASIA, AMERICA, Australia-Pacific (AUSPAC)).
DEMAND	Second component of a PCA of the 14 single scores which constitute GGGI. The ratio of female to male wages for similar work (WAGEEQUALITY), the ratio of healthy male to female lifespans (LIFEEXPECTANCY), and the ratio of a female compared to a male head of state over the last 50 years (FEMHEADSTATE), as well as estimated earned income (EARNINCOME) load significantly on this component (factor loading >.3).
GENDERCULTURE	Value for cultural gender egalitarianism based on the GLOBE project (Source: House et al. 2004).
GDPCAP	Natural logarithm of GDP per capita (World Bank data code: NY.GDP.PCAP.CD).
GGGI	Index value of the Global Gender Gap Index (Source: World Economic Forum).
GOVCODE	Indicator variable equal to 1 if the country's non state-owned firms are subject to voluntary gender regulation in good governance codes, and 0 otherwise (Source: Terjesen et al. 2015; additional hand-collection).
GQBOARD	Average ratio of female board members (Asset4 code: CGBSO17V) based on a minimum of 10 firm observations per country-year.
INVPROT	Index value of the minority investor protection index compiled by the World Bank based on three subindices (i.e., (1) Extent of Disclosure index, (2) Extent of Director Liability index, and (3) Ease of Shareholder suit index).
QUOTA	Indicator variable equal to 1 if the country's non state-owned firms are subject to a mandatory gender quota on the corporate board, and 0 otherwise (Source: Terjesen et al. 2015; additional hand-collection).
LEGOR	Indicator variable equal to 1 if the country follows a certain legal tradition (e.g., French (LEGOR_FR), UK (LEGOR_UK), Germanic (LEGOR_GE), Scandinavian (LEGOR_SC)) (Source: LaPorta et al. 1998; Klerman et al. 2011).
SUPPLY	First component of a Principal Component Analysis (PCA) of the 14 single scores which constitute GGGI. The ratio of female over male senior officials and managers (MANAGERS), the ratio of female over male professional and technical workers (PROFESSIONALS), the ratio of female over male literacy (LITERACY), the ratios of female over male enrolment in secondary (SECONDARY) and tertiary (TERTIARY) education, as well as estimated earned income (EARNINCOME) load significantly on this component (factor loading >.3).

CHAPTER 4:
BOARD GENDER DIVERSITY AND ITS EFFECTS ON CAPITAL MARKETS – NEW INTERNATIONAL EVIDENCE

Abstract

We analyze investors' perception and long-term effects of board gender diversity on firms' capital market performance in an international setting. Our results, controlling for the endogenous nature of board appointments, indicate that female board representation neither improves nor reduces firms' long-term stock performance. Thus, investors seem to perceive female and male board members as being equivalent in the long-term and, on average, do not base their investment decisions on directors' gender. Hence, we argue that it is imperative to go beyond the conventional thinking in terms of the business case for gender diversity and broaden the perspective also in order to incorporate societal and ethical aspects in the strive to board gender equality. Even more so, our results show that it does not entail reduced shareholder value, which the literature on mandatory gender quotas commonly seems to suggest.

This part of the thesis is a joint project with Thomas R. Loy. This paper is currently under review at *Business & Society* (3rd round).

Acknowledgements: We gratefully acknowledge comments and suggestions by Marcus Bravidor, Douglas Cumming, Benedikt Downar, Emilie Feldman, Rolf Uwe Fülbier, Gerard George, Maria Goranova, Aparna Joshi, David Reeb, Jan Seitz, delegates at the 38th EAA Annual Congress in Maastricht, delegates at the 78th Annual Congress of the German Academic Association of Business Administration in Munich, and seminar participants at University of Bayreuth. We are especially grateful to the World Economic Forum for providing historical data on the Global Gender Gap Index. We also thank Isabell Keller, Sophie Kobes, Marie-Thérèse Meyer and Johanna Paraknewitz for excellent research assistance.

1 Introduction

Female board representation has been extensively examined in prior research with respect to its business case. As such, a majority of studies focuses on business-related arguments, such as improved corporate governance (e.g., Fondas 2000) as well as financial performance (for meta-analysis, c.f. Post and Byron 2015; Pletzer et al. 2015). However, while most prior studies find a positive association of female board representation with corporate governance activities (e.g., Adams and Ferreira 2009; Nielsen and Huse 2010a, 2010b), there are inconclusive results with respect to its impact on firms' financial performance. In this context, proponents argue that gender diversity improves the quality of board discussions, attributable to more creativity (e.g., Nemeth 1986) and a wider range of perspectives (Hillman et al. 2007). Critics point to more conflicts between board members which arise from internal divisions and social categorization tendencies (Tajfel 1979). As such, it is not particularly surprising that some studies find a positive association (e.g., Dezsö and Ross 2012; Erhardt et al. 2003; Krishnan and Park 2005; Singh et al. 2001), while others conclude with the opposite or no significant effects (e.g., Adams and Ferreira 2009; Carter et al. 2010; Rose 2007; Shrader et al. 1997).

Whereas most empirical research focuses on accounting performance, there is limited evidence on the association of board gender diversity with market performance. To this day, there are only a few studies analyzing short-term market effects of female board appointments, which likewise exhibit mixed results (e.g., Farrell and Hersch 2005; Kang et al. 2010; Lee and James 2007; Schmid and Urban 2016). To our surprise, previous studies have neglected the long-term market performance attributable to female board representation.

Explicitly, taking a longer-term perspective is valuable for multiple reasons. First, short-term market reactions proxy for the expected impact of female appointees on shareholder value. We are more interested in the actual association of gender diverse boards with long-term changes in shareholder wealth. Second, new board appointees require some time to enact organizational changes and shifts in firms' investment patterns (Pan et al. 2016). The market also undergoes an adjustment period in which it evaluates the appointees' quality (Pan et al. 2015). Due to this, there ought to be longer-term effects of female board representation on a firm's stock performance. Third, a firm's focus is growth

and going-concern and, therefore, to increase shareholder value. In connection, there are calls for increased board gender diversity from a wide variety of stakeholders, such as institutional investors (Byoun et al. 2016; Coffey and Fryxell 1991), stock exchanges (Terjesen et al. 2016), as well as regulators, such as the Securities and Exchange Commission (SEC) (SEC 2009). This external pressure might result in firms catering to the market demand for board gender diversity, or being punished if they fail to do so. In sum, there is a lack of evidence on the "relationship between board diversity and long-term stock price performance which is the 'gold standard' measure of shareholder value" (Rhode and Packel 2014, p. 391).

Therefore, it seems important to analyze the effect of female corporate board participation and long-term changes in shareholders wealth. As a result, we attempt to answer the following two research questions (RQ):

RQ 1*: Is female board participation associated with systematic differences in the long-term development of investors' wealth in a global setting?*

RQ 2*: Is female board participation associated with systematic differences in investors' perception of firms which do (not) cater to the market demand for board gender diversity?*

Focusing on our two research questions, we provide several important contributions to current research. First, our study contributes to the social and economic debate regarding female board representation by analyzing how stakeholders (e.g., investors) perceive female board members and their contribution to firms' profitability. Second, we investigate whether aforementioned societal pressures result in firms catering to these expectations or, if they fail to 'comply', whether they are punished by investors. Third, corporate board composition and its association with performance is a profoundly endogenously issue (Hermalin and Weisbach 2003), which is largely ignored by the majority of previous studies. We explicitly address this gap in the literature by employing Propensity Score Matching (PSM). Fourth, prior empirical studies have mainly focused on single countries (i.e., predominantly the U.S.). However, prior cross-country research shows that female board representation differs significantly between countries (e.g., Grosvold and Brammer 2011; Grosvold et al. 2016; Loy and Rupertus 2018a; Terjesen and Singh 2008) and that there are still global differences in overall gender equality (WEF 2017). Hence, we test our research questions in an international setting.

Based on our analyses, we provide some interesting new insights into the relationship of female board participation and long-term market performance. Across a comparable set of firms, our results indicate no significant differences in annual stock returns. Moreover, we fail to document significantly reduced stock returns for firms that deviate from the expected ratio of female board membership. While female board appointments may have the reported short-term effects (e.g., Kang et al. 2010; Lee and James 2007; Schmid and Urban 2016), it seems that the market corrects this (negative) mispricing over the long run. In sum, our results do not suggest that improvements in gender equality are not a worthwhile undertaking. From the investors' perspective, high-quality board members seem equivalent in the long term, regardless of their gender. Hence, implementing unbiased hiring practices are a societal imperative for reasons of fairness, equality, and participation (e.g., Grosvold et al. 2007; Singh et al. 2015).

We structure our paper as follows. First, we develop our theoretical framework, discuss prior literature, and develop our hypotheses. Second, we explain our empirical methods. Third, we present the data and descriptive statistics. Fourth, we exhibit and discuss our multivariate results. Fifth, we challenge our findings through a range of robustness checks and discuss the implications of our results. The final section concludes.

2 Theoretical Framework, Prior Literature and Hypothesis Development

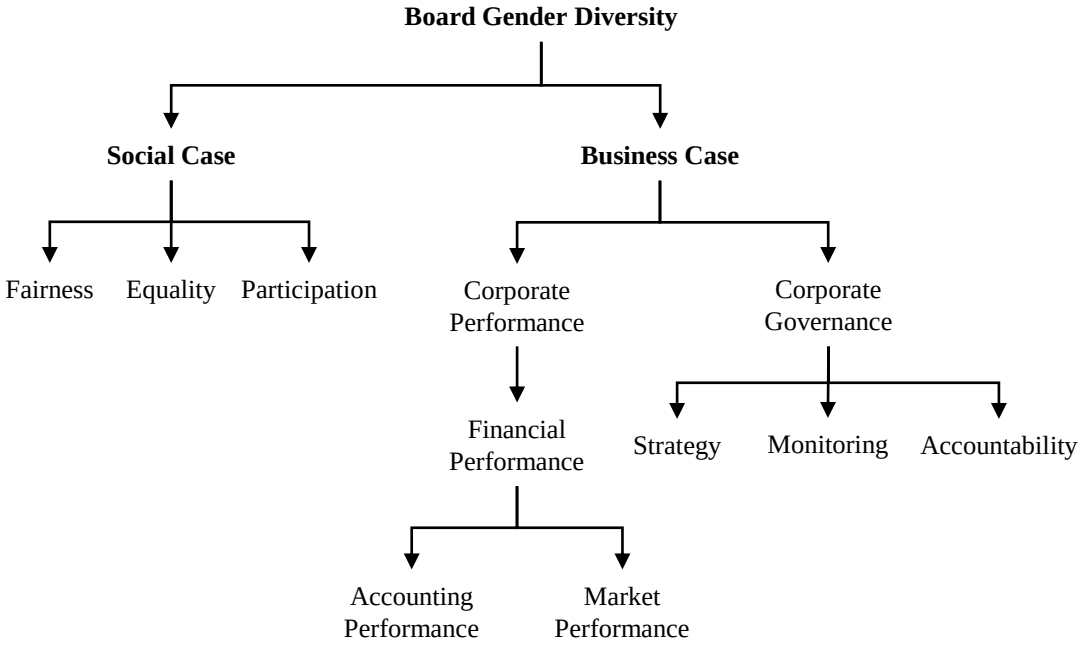
2.1 The Social and Business Case of Board Gender Diversity

Internationally, board gender equality and the lack of female board representation has increasingly become the focus of political and societal debates in recent years (e.g., Grosvold 2011; Pande and Ford 2011; Singh et al. 2015; Terjesen et al. 2016). Despite international efforts to increase female board representation, there still remain huge discrepancies in the proportion of female board members across and within countries (e.g., Loy and Rupertus 2018a; Terjesen and Singh 2008).

Thus, men still dominate corporate boards and there seem to exist still significant barriers for women to reach the top of corporate management (i.e., the 'glass ceiling effect', e.g., Arfken et al. 2004). Whereas several studies focus on the causes of these major international discrepancies in the number of women on corporate boards (e.g., Grosvold 2011;

Grosvold and Brammer 2011; Grosvold et al. 2016; Loy and Rupertus 2018a; Terjesen and Singh 2008), others focus on firm-level predictors for female board representation (e.g., Farrell and Hersch 2005; Hillman et al. 2007; Terjesen et al. 2009). In recent years, scientific studies put an increased focus on the motives for greater board gender diversity and rely upon social as well as business-related theories and lines of argument (Campbell and Minguez Vera 2010; Grosvold et al. 2016; Kilgour 2013; Singh et al. 2015).

Figure 9
The Social and Business Case of Board Gender Diversity



With respect to the former, social arguments for greater board gender diversity include, among others, fairness, equal opportunities, and participation (e.g., Brammer et al. 2007; Rhode and Packel 2014; Singh et al. 2015) (Figure 9). As today's women are more highly educated and participate in considerably more professional and technical occupations than their mothers and grandmothers (WEF 2016), continuing a pattern of unequal opportunities implies a great waste of talent and, thus, societal dead-weight-loss (Radjavi 2012). In summary, proponents of the current discussion with respect to increasing the number of female board members argue that today's society overall benefits from gender diverse corporate boards. In line with this, firms have the distinct opportunity to send a positive signal to their share- and stakeholders by appointing women (e.g., Grosvold et al. 2007).

The business case of board gender diversity mostly relies on the assumption that female board representation improves organizational processes and performance (Rhode and Packel 2014). This includes corporate governance outcomes as well as firms' financial performance (Grosvold et al. 2016). On the one hand, firm performance entails three dimensions: (1) Long-term financial (i.e., accounting) performance, (2) long-term market performance, as well as (3) short-term market reactions. On the other hand, in the context of corporate governance boards of directors have three central functions: (1) Shaping strategy, (2) monitoring top executives, as well as (3) enhancing accountability (Fondas 2000; Post and Byron 2015). Among others, boards are responsible for several tasks which are crucial for the corporation. For instance, the board selects and replaces the chief executive officer (CEO), provides advice and counsel to top management, represents the interests of shareholders, as well as monitors and controls management and company performance. If boards efficiently fulfil these tasks, they can (positively) affect a company's performance. More specifically, upper echelons theory (UET) establishes a framework in which board members' unobservable cognitive base values (e.g., limited field of vision, selective perception, and means of acquiring and interpreting new information) or observable characteristics (e.g., age, gender, or career experience) determine a board's performance in its primary tasks, its organizational practices, and approaches to strategy formulation (Hambrick 2007; Hambrick and Mason 1984; Zahra and Pearce 1989). Based on these arguments, a large body of research suggests that female board members determine the boards' activities and thus, may positively (or negatively) affect a company's (financial) performance as well as corporate governance outcomes (Post and Byron 2015).

Proponents argue that women bring useful female leadership qualities and skills to the boards which improve corporations' decision-making processes through fresh and well-informed views on market and environmental issues (Boulouta 2013), risk awareness, as well as less radical and less overconfident decision making (Chen et al. 2016; Huang and Kisgen 2013; Jianakoplos and Bernasek 1998). Moreover, gender diverse boards are more likely to engage in high-quality analysis and are less likely to take extreme positions (e.g., Dobbin and Jung 2011; Rhode and Packel 2014). By appointing more women to boards, the organization sends a signal that it offers fair opportunities for career progress for current and future female employees (Hillman et al. 2007). If firms do not provide gender-neutral access into corporate boards, they will lack female skills and perspectives (Holton

2000; Terjesen and Singh 2008). Hence, it enjoys and retains access to a broader pool of qualified present and potential employees on all organizational levels. Critics imply that increased female board representation may generate more diametrically opposed opinions and critical questions during board meetings (Lau and Murnighan 1998). Thus, decision-making is less effective and more time-consuming (Miller et al. 1998). Moreover, Tajfel (1979) suggests that social categorization tendencies lead to gender salience and thereby more stereotypes within groups (Abrams et al. 1990). This can hinder functional team processes and possibly lead to increased inconsistencies in communication and cooperation (van Knippenberg et al. 2004). Hence, conflicts between board members might outweigh the benefits of additional perspectives.

2.2 Prior Literature

To this day, extensive research shows that increased gender diversity has a positive impact on board activities. For instance, Adams and Ferreira (2009) find that female board members have better attendance rates and are more inclined to join monitoring committees. Nielsen and Huse (2010a, 2010b) find that female directors contribute positively to strategic board decisions, board effectiveness, and that boards with higher ratios of women have more board development evaluations and programs. Moreover, boards with a higher proportion of women hold more board meetings, are more likely to replace a CEO when stock performance is poor, and are positively associated with better board monitoring (e.g., Adams and Ferreira 2009). Overall, prior studies indicate a positive association of board gender diversity with the effectiveness of corporate boards.

Research assumes that if board gender diversity affects corporate boards' efficiency, there should be an effect on e.g., financial as well as market performance (Dobbin and Jung 2011). So far, the literature has extensively focused on the association of female board membership with firms' financial accounting performance. One strand of research finds a positive relationship (e.g., Dezsö and Ross 2012; Erhardt et al. 2003; Krishnan and Park 2005; Singh et al. 2001). Yet, another strand finds the opposite or no significant relationship (e.g., Adams and Ferreira 2009; Carter et al. 2010; Rose 2007; Shrader et al. 1997). Unsurprisingly, especially more recent studies argue that the link between firms' financial performance and female board representation is complex and, above all, indirect (Forbes and Milliken 1999; Galbreath 2018).

With respect to the link of gender diversity and equity market performance, there is only a handful of studies analyzing short-term market reactions to the appointment of female board members. Whereas Kang et al. (2010) find positive short-term investor reactions to female board appointments in Singapore, Schmid and Urban (2016) take a more nuanced stance. They provide evidence that stock markets exhibit more negative short-term reactions to deaths of female board members, especially in countries with low rates of female board representation. However, the authors state that this effect is largely clustered in countries with a stronger 'glass ceiling' and, thus, potentially attributable to more rigorous screening of female directors. On the contrary, Lee and James (2007) find negative announcement returns to female CEO appointments. And finally, Farrell and Hersch (2005) find a positive association between firm performance and the likelihood of appointing women to the board, but they do not find support for an associated market reaction. Thus, also prior studies analyzing short-term market effects find mixed results.

Nevertheless, it is doubtful that a focus on short-term market effects can effectively measure the impact of board gender diversity on investors' perception. Quite surprisingly, previous studies have neglected to investigate the long-term change in shareholder wealth attributable to female board representation, which is "the 'gold standard' measure of shareholder value" (Rhode and Packel 2014, p. 391). Our focus on the longer-term perspective is valuable for two reasons. First, short-term market reactions proxy for the expected impact of female appointees on shareholder value. Thus, it cannot capture the actual association. Second, new board appointees require some time to enact organizational changes and changes in firms' investment patterns (Pan et al. 2016). Indicative of this, Pan et al. (2015) find evidence for an almost linear decline in equity volatility over a CEO's first year of tenure. On average, stock volatility is highest at the appointment date, which likely biases short-term appointment returns. In the following, the market undergoes an adjustment period during which it evaluates the quality of appointees, beyond a potential stereotypical bias. As a result, looking at the longer-term perspective seems to be a more convincing way to examine investors' perception of female board representation and its actual effects.

2.3 Hypothesis Development

Role congruity theory in psychology posits that women are regarded as less favorable candidates for leadership roles compared to men as a result of a perceived incongruity between female roles and leadership roles. Eagly and Karau (2002) advance existing theories regarding prejudice and combine them with perceptions of managerial roles. More specifically, prejudice against women may arise when social perceivers have stereotypes about a social group that is incongruent with well-known and expected attributes of roles of another group. Once a negatively stereotyped person rises to the top of the corporation and, hence, an incongruent social role, this inconsistency diminishes the appreciation this person receives from others (Eagly and Karau 2002). In conclusion, role congruity theory explains that stereotypes of female gender roles might prescribe less leadership qualities to women than to men.

Following this theory, one could argue that the effect of female board representation on capital market performance occurs mainly through societal stereotyping that may influence investors' evaluation of a firms' potential to generate future earnings (Post and Byron 2015). If individuals that deviate from social expectations occupy specific positions in the firm, investors may react negatively to firms with more female board members (e.g., Dobbin and Jung 2011). As international evidence shows, there still is a relatively low number of women on corporate boards (e.g., Grosvold and Brammer 2011; Grosvold et al. 2016; Loy and Rupertus 2018a; Terjesen and Singh 2008). Hence, this probably reinforces stereotypes about female (dis)qualifications for such positions. A large body of literature in organization theory shows that especially men believe that leadership positions are associated with masculine qualities (e.g., Gupta et al. 2009; Powell and Butterfield 2002; Schein and Mueller 1992). Unsurprisingly, they view female appointees with considerable skepticism (Dobbin and Jung 2011; Lee and James 2007) and point out the detrimental effects of imposing board gender diversity through regulation (Ahern and Dittmar 2012). Moreover, investors may not be accustomed to women on corporate boards for the reason, which on the surface might seem more legitimate, that they lack business experience for such positions (Dobbin and Jung 2011). However, such negative preconceptions towards female board appointees would likely culminate in negative short-window stock reactions. As time progresses, investors may update their beliefs and revise their negative stereotypes to reflect the positive effects of newly implemented strategies. Therefore, investors may continue investing to the extent that they are satisfied

with the female board appointees. On the contrary, they will discontinue or lower future investments if they are dissatisfied with the board's direction (Lee and James 2007). Cumulatively, we state our first hypothesis in positive form, but the issue remains an empirical question:

H1: *Firms with gender diverse corporate boards exhibit improved long-term equity capital market performance compared to their peers without women on their boards.*

Additionally, corporations are increasingly attentive to the desires of a wide range of stakeholders (e.g., national governments, politicians, employees) as well as shareholders (e.g., sovereign wealth funds, pension funds, or large institutional investors). On the one hand, there are numerous non-profit initiatives to promote female board representation in basically every highly-developed capital market. They aim to convince businesses to implement gender fair corporate boards, even if this is not mandatory by law – or go beyond the quota. Additionally, board gender diversity is regarded as a key pillar of good governance and fixated in numerous voluntary corporate governance codes, globally (Loy and Rupertus 2018a; Terjesen et al. 2015). The SEC even issued a rule that listed companies must disclose their views on board diversity (SEC 2009). On the other hand, there is evidence that investors pay attention to board structures (e.g., Gillan and Starks 2000; Yermack 2006) and make board diversity, individually and collectively, a higher priority in investment decisions (Dobbin and Jung 2011; Rhode and Packel 2014). Moreover, there is an increasing number of (institutional) investors who actively push firms to increase their board gender diversity (Byoun et al. 2016; Coffey and Fryxell 1991; Fondas 2000). Such groups can leverage their exposed position to influence internal decisions in order to advance female board representation among companies in which they hold significant stakes (Dobbin and Jung 2011; Rhode and Packel 2014).

In conclusion, we argue that equity markets form their own expectations about the ratio of female board members. If key players in equity markets demand a higher proportion of female board members, firms may face significant pressure to increase female board representation (Dobbin and Jung 2011; Rhode and Packel 2014; Terjesen et al. 2009). While firms' fundamentals might not have changed too much over recent years, they still might feel inclined to cater to a market demand for female board representation or run the risk of being punished by investors over the long-term (Ghosh et al. 2016). As a result,

we argue that if firms do not meet investors' expectations about gender diverse boards, they might be punished by a decline in long-term stock returns. Thus, we state our second hypothesis as follows:

H2: *Firms exhibit impaired long-term stock market performance if the observable ratio of female board representation is below market expectations.*

3 Empricial Models

3.1 Propensity Score Matching

An ideal empirical model would establish and test a causal relationship between explanatory and explained variables. A necessary condition is that explanatory variables are exogenous, and not prone to self-selection. In our context, corporate board composition is profoundly endogenous (Adams and Ferreira 2009; Hermalin and Weisbach 2003). The pure existence of corporate boards is exogenous – based on stock exchange requirements or incorporation laws, both well outside individual firms' control. Nevertheless, their composition is largely endogenous, with some notable exceptions like mandatory gender quotas or co-determination with mandatory employee board representation (i.e., particularly in the German case). As such, detecting a specific effect of female board representation on (market) performance is not trivial.

There are only a handful of prior studies, which appropriately engage endogeneity. While one meta-analysis (Pletzer et al. 2015) does not mention the issue at all, Post and Byron (2015) specifically name two-stage models as a potential remedy to be applied in future research. We implement this through a quasi-experimental approach, novel to this literature. PSM builds on identifying potentially relevant covariates, based on previous empirical findings and economic theory, which influence selection into the treatment group (Rosenbaum and Rubin 1983). The control group is formed by sampling a large number of potential control firms for those which exhibit highly similar covariates to the treatment group. Incidentally, this answers a recent call for future research which deems "board selection processes (…) a fruitful avenue for future analysis" (Post and Byron 2015, p. 1562). Our PSM approach is based on the board gender diversity prediction model by Hillman et al. (2007) and Gul et al. (2011). To control for country-specific regulations and institutional factors, which likely also affect the ratio of female board

members, we estimate the conditional probability of receiving treatment on a country-by-country basis (Equation 8):

$$\begin{aligned} GENDER_{i,t} = {} & \alpha_0 + \beta_1 ROA_{i,t} + \beta_2 SIZE_{i,t} + \beta_3 BOUTSIDE_{i,t} + \beta_4 RISK_{i,t} \\ & + \beta_5 TD_{i,t} + \beta_6 AGE_{i,t} + \beta_7 SGRWOTH_{i,t} + \beta_8 LEV_{i,t} \\ & + \beta_j \Sigma_j IND_{i,t} + \beta_k \Sigma_k YEAR_{i,t} + \varepsilon \end{aligned} \quad (8)$$

Return on Assets (ROA). Hermalin and Weisbach (1988) examine determinants of changes among corporate directors. They hypothesize that firm performance has the utmost impact. If a firm's financial performance is poor, incumbent directors are more likely to be removed and replaced by new hires. While the incumbent directors are likely male, new appointees can be of both genders. A number of studies finds a positive relationship between ROA and board gender diversity (e.g., Carter et al. 2010; Erhardt et al. 2003; Post and Byron 2015). On the contrary, others do not find significant associations (e.g., Dobbin and Jung 2011). Therefore, we refrain from a directional prediction.

Firm Size (SIZE). Firm size covers a number of unobservable firm characteristics, such as public visibility and political costs. As such, Adams and Ferreira (2009) argue that larger firms exhibit increased demand for board diversity since they are subject to more public scrutiny and pressure to meet social expectations. Increased variety and a larger number of stakeholders, more common for larger firms, contribute to this social pressure (Hillman et al. 2007). Prior studies establish a positive relationship between the percentage of women on corporate boards and firm size (e.g., Carter et al. 2010). Therefore, we expect that firm size is likely positively associated with the probability of female board representation.

Board Outsider (BOUTSIDE). Firms form interorganizational networks through their top managers' outside directorships and other firms' directors, to deal with environmental uncertainties as well as to obtain valuable information and important resources (Hillman et al. 2007; Pfeffer 1972). Interlocking directorates, thus, may provide additional information about a supply of female directors which could lead to a diffusion of organizational practices, such as gender diversity (Hillman et al. 2007). Moreover, appointing additional outside directors allows firms to add female perspectives to the board without necessarily having to replace experienced and suitable male directors. Therefore, we expect a positive association between the number of outside directors and female board representation.

Risk (RISK). A firm's risk profile seems to be one of the most important and robust factors to determine the proportion of women on boards of directors (Adams and Ferreira 2009; Farrell and Hersch 2005). In both studies, the authors present a statistically significant, negative relationship between gender diversity and firm risk. These findings are in line with results by Chen et al. (2016) and Jianakoplos and Bernasek (1998), who determine that women are more risk averse with regard to financial and research & development (R&D) investment decisions. On the contrary, Adams and Funk (2012) exhibit that female top executives are significantly less risk averse than their male counterparts. Therefore, we do not predict a sign.

Total Diversification (TD). We include the total diversification measure of Palepu (1985) as an indicator for firm strategy. Hillman et al. (2007) argue that a firm, which only operates in a single business, is environmentally more path dependent. On the contrary, multi-business corporations are inherently more prone to a broader set of perspectives and ties to more (diverse) stakeholders. As a result, we expect increased female board representation for more diversified firms.

In accordance with Gul et al. (2011) and Hillman et al. (2007), we further include controls for firm age (AGE), sales growth (SGROWTH), and the debt-to-equity ratio (LEV). Additionally, we include industry-fixed effects (IND), because there is ample evidence which suggests that industry is a significant explanatory factor of female board representation (e.g., Cumming et al. 2015), as well as year-fixed effects (YEAR) to control for a potential trend of increased female board representation in recent years (Loy and Rupertus 2018a). To control for outliers, the values of all non-truncated variables are winsorized at the 1st and 99th percentile.

Our dependent variable GENDER is a dummy variable equal to 1 if a firm has at least one female board member, and 0 otherwise. We match each firm-year from the treatment group (GENDER = 1) in a given country (e.g., Spain) with another firm-year of the control group (GENDER = 0) from the same country with the closest propensity score, without replacement. Conceptually, we expect that matched firms only differ with respect to the treatment (i.e., board gender diversity).

3.2 Value Relevance Design

To further test our first hypothesis – the association of female board membership with long-term development of shareholder wealth – we employ a multivariate value relevance (VR) design. As such, we pool all country data sets and run the following regression:

$$RET_{i,t} = \alpha_0 + \beta_1 GENDER_{i,t} + \beta_2 EPS_{i,t} + \beta_3 \Delta EPS_{i,t} + \beta_j \Sigma_j IND_{i,t} + \beta_k \Sigma_k YEAR_{i,t} + \varepsilon \quad (9)$$

$$RET_{i,t} = \alpha_0 + \beta_1 GENDERMATCH_{i,t} + \beta_2 EPS_{i,t} + \beta_3 \Delta EPS_{i,t} + \beta_j \Sigma_j IND_{i,t} + \beta_k \Sigma_k YEAR_{i,t} + \varepsilon \quad (10)$$

The dependent variable is annual stock return corrected for possible dividend payments (RET) for the period nine months before through three months after fiscal year-end. This period corresponds to the disclosure of annual financial statements and the timing of annual shareholders' meetings at which new directors are voted into office. EPS is computed as net income before extraordinary items scaled by the number of shares outstanding. ΔEPS is the year-to-year change in EPS scaled by the stock price at the beginning of the fiscal year. Most importantly, we examine the incremental effect of information about female board representation. In the first regression (Equation 9), we build on the full sample and include the unmatched GENDER variable to estimate its predictive ability. Attributable to aforementioned endogeneity concerns, we concentrate on the matched sample (GENDERMATCH) in a second analysis (Equation 10). Both regressions include controls for industry- (IND) and year- (YEAR) fixed effects. To control for outliers, values of all non-truncated (i.e., non-binary) variables are winsorized at the 1st and 99th percentile, as well.

3.3 Gender Diversity Expectation Model

To test our second hypothesis, we extend our value relevance design. We analyze whether deviations from the expected ratio of female board members affect shareholders long-term investment decisions (Hypothesis 2). Therefore, we include ABSGENDIFF as the absolute difference between the observed proportion of female board members in a given firm and the expected ratio of female board representation (Equation 11):

$$RET_{i,t} = \alpha_0 + \beta_1 EPS_{i,t} + \beta_2 \Delta EPS_{i,t} + \beta_3 ABSGENDIFF_{i,t} + \beta_4 NEGGAP_{i,t} + \beta_5 ABSGENDIFF \times NEGGAP_{i,t} + \beta_j \Sigma_j IND_{i,t} + \beta_k \Sigma_k YEAR_{i,t} + \varepsilon \quad (11)$$

Prior research suggests that firms in certain industries are more likely to appoint women to leadership positions (e.g., Cumming et al. 2015). Thus, to determine the expected ratio of female board members, we use the average ratio of female board representation across all firms in a given industry-country cross-section. Moreover, we include an indicator variable equal to one if a firm's observed proportion of female board members downwardly deviates from the expected country-industry-based ratio of women on the board (NEGGAP). As we are primarily interested in the incremental effect of a downward deviation from expectations, we include the interaction term of both variables (ABSGENDIFF × NEGGAP). Hence, the baseline effect ABSGENDIFF measures the effect of a positive deviation. Finally, we include controls for industry- (IND) and year- (YEAR) fixed effects.

4 Sample Selection and Descriptive Statistics

4.1 Sample Selection

Initially, we consider all firms included in the Thomson Reuters' Asset4 database for the period between 2008 and 2014. Main variables for all our investigations are board characteristics with respect to the total number of board members and the percentage of women on the board. We limit our sample to firms from countries with a sufficient number of firm-years with available board characteristics and financial data from Worldscope to carry out our PSM estimations. Moreover, we eliminate firm-year observations with missing stock returns. Our final sample consists of 8,872 firm-year observations from 13 countries.

Table 32
Geographical Sample Distribution

Country	Total	2008	2009	2010	2011	2012	2013	2014
Australia	532	40	51	78	89	90	88	96
Brazil	100	4	9	17	13	28	11	18
Canada	524	67	84	85	75	77	65	71
France	291	66	68	73	20	22	20	22
Germany	302	49	56	54	34	40	31	38
India	161	6	14	19	28	33	27	34
Italy	109	26	31	31	1	8	5	7
Japan	2,272	319	321	345	323	325	307	332
Singapore	116	24	25	27	10	10	10	10
Spain	128	25	26	26	7	16	14	14
Switzerland	218	34	33	40	26	29	29	27
U.K.	427	72	75	78	50	51	50	51
U.S.	3.692	411	540	598	595	586	452	510
Σ	8,872	1,143	1,333	1,471	1,271	1,315	1,109	1,230

Table 32 presents the geographical sample distribution.

We present our pooled sample by country in Table 32 and our unmatched (Column a) as well as matched (Column b) sample by year (Panel A), country (Panel B), and industry (Panel C) in Table 33. There is significant variation in the number of firm-years across countries. A majority of observations originate from Australia, Canada, Japan, the U.K., and the U.S. There are also unambiguous differences between the treatment (GENDER = 1) and comparison group (GENDER = 0). On the one hand, the majority of U.S. firms have at least one woman on their boards of directors despite that there are no legislative plans to implement a mandatory gender quota. France, Germany, the U.K., and Canada present similar patterns. On the other hand, most firms from Japan do not exhibit gender diverse boards.

Table 33
Unmatched and Matched Sample Distribution

	(a) Unmatched Sample					(b) Matched Sample				
		GENDER = 0		GENDER = 1			GENDER-MATCH = 0		GENDER-MATCH = 1	
	Total	N	%	N	%	Total	N	%	N	%
Panel A: By Year										
2008	1,143	488	14.78	655	11.76	303	148	13.11	155	13.73
2009	1,333	536	16.24	797	14.31	374	185	16.39	189	16.74
2010	1,471	559	16.93	912	16.37	410	212	18.78	198	17.54
2011	1,271	470	14.24	801	14.38	344	168	14.88	176	15.59
2012	1,315	464	14.06	851	15.28	314	161	14.26	153	13.55
2013	1,109	382	11.57	727	13.05	222	108	9.57	114	10.10
2014	1,230	402	12.18	828	14.86	291	147	13.02	144	12.75
Σ	8,872	3,301		5,571		2,258	1,129		1,129	
Panel B: By Country										
Australia	532	132	4.00	400	7.18	202	101	8.95	101	8.95
Brazil	100	56	1.70	44	0.79	34	17	1.51	17	1.51
Canada	524	98	2.97	426	7.65	84	42	3.72	42	3.72
France	291	38	1.15	253	4.54	56	28	2.48	28	2.48
Germany	302	73	2.21	229	4.11	102	51	4.52	51	4.52
India	161	77	2.33	84	1.51	68	34	3.01	34	3.01
Italy	109	50	1.51	59	1.06	36	18	1.59	18	1.59
Japan	2,272	2,024	61.31	248	4.45	488	244	21.61	244	21.61
Singapore	116	51	1.54	65	1.17	58	29	2.57	29	2.57
Spain	128	25	0.76	103	1.85	26	13	1.15	13	1.15
Switzerland	218	77	2.33	141	2.53	50	25	2.21	25	2.21
U.K.	427	119	3.60	308	5.53	158	79	7.00	79	7.00
U.S.	3,692	481	14.57	3,211	57.64	896	448	39.68	448	39.68
Σ	8,872	3,301		5,571		2,258	1,129		1,129	
Panel C: By Industry										
NonDur	569	192	5.82	377	6.77	126	70	6.20	56	4.96
Dur	421	259	7.85	162	2.91	92	50	4.43	42	3.72
Manuf	1,200	583	17.66	617	11.08	259	131	11.60	128	11.34
Enrgy	359	132	4.00	227	4.07	136	71	6.29	65	5.76
Chems	502	212	6.42	290	5.21	124	69	6.11	55	4.87
BusEq	921	356	10.78	565	10.14	300	145	12.84	155	13.73
Telcm	291	90	2.73	201	3.61	62	31	2.75	31	2.75
Utils	555	141	4.27	414	7.43	55	26	2.30	29	2.57
Shops	757	229	6.94	528	9.48	194	101	8.95	93	8.24
Hlth	340	81	2.45	259	4.65	46	20	1.77	26	2.30
Money	1,602	408	12.36	1,194	21.43	325	162	14.35	163	14.44
Other	1,355	618	18.72	737	13.23	539	253	22.41	286	25.33
Σ	8,872	3,301		5,571		2,258	1,129		1,129	

Table 33 presents detailed sample distributions for the unmatched and matched sample by year (Panel A), country (Panel B), and industry (Panel C).

4.2 Descriptive Statistics

For the pooled sample, average female board representation (GENQUOT) is 10.9 percentage points with a median of 10.0 (Table 34). Their mean (median) size (natural logarithm of total assets) is 16.1 (15.9) and the average (median) age, which corresponds to the period since the date of incorporation, is 40 (45) years. Furthermore, outside directors occupy on average (median) 6 (7) seats on the corporate boards and the mean (median) firm-specific risk is 0.036 (0.026).

Research into the effects of board gender diversity is subject to severe endogeneity problems (Hermalin and Weisbach 2003) because firms with gender diverse boards could be systematically different from all-male board firms (Adams and Ferreira 2009; Hillman et al. 2007; Lai et al. 2017). Comparing our gender diverse (N=5,571) and non-gender diverse (N=3,301) sub samples, substantial and mostly significant differences along a number of dimensions are quite visible (cf. Table 37).

Compared to their counterparts, firms with female board presence exhibit on average (median) higher stock returns of 4.8 (8.1) percent, are larger, younger and have more outside directors. Furthermore, they exhibit a lower risk profile. Generally, these descriptive statistics are in line with prior research (e.g., Hillman et al. 2007).

Table 34
Descriptive Statistics – Pooled Sample

Variables	Mean	25 %	Median	75 %	SD	N
GENQUOT	0.109	0.000	0.100	0.182	0.571	8,872
GENDER	0.628	0.000	1.000	1.000	1.000	8,872
GENDERMATCH	0.500	0.000	0.500	1.000	0.500	2,258
RET	0.119	-0.210	0.058	0.345	0.512	8,872
ROA	0.040	0.010	0.034	0.070	0.061	8,872
SIZE	16.14	15.03	15.92	17.08	1.58	8,872
BOUTSIDE (ln)	1.839	1.609	1.946	2.197	0.525	8,872
BOUTSIDE (#)	6	5	7	9	2	8,872

Table 34 – continued
Descriptive Statistics – Pooled Sample

Variables	Mean	25 %	Median	75 %	SD	N
RISK	0.036	0.015	0.026	0.044	0.035	8,872
TD	0.877	0.553	0.871	1.204	0.464	8,872
AGE (ln)	3.688	3.091	3.807	4.382	0.838	8,872
AGE (#years)	40	22	45	80	2	8,872
SGROWTH	0.079	-0.005	0.050	0.120	0.203	8,872
LEV	3.219	0.811	1.517	3.043	5.644	8,872
EPS	0.052	0.031	0.059	0.087	0.096	8,872
ΔEPS	-0.008	-0.023	-0.001	0.011	0.797	8,872
GENDIFF	0.000	-3.679	-0.610	3.976	7.699	8,872
ABSGENDIFF	5.485	1.060	3.750	8.334	5.401	8,872
NEGGAP	0.603	0.000	1.000	1.000	0.489	8,872
ABSGENDIFF × NEGGAP	2.743	0.000	0.610	3.679	4.299	8,872

Table 34 presents summary descriptive statistics for all firm-year observations of the pooled sample for the period 2008 through 2014. GENQUOT is the observed percentage of women on the boards of directors. GENDER is an indicator variable equal to 1 if a firm is included into the treatment group (has at least one female board member), 0 otherwise. GENDERMATCH is an indicator variable equal to 1 if a firm is included into the treatment group after applying the PSM model, 0 otherwise. RET is stock returns. ROA is return on assets. SIZE is the natural logarithm of total assets. BOUTSIDE is the natural logarithm of outside directors. RISK is the standard deviation of cash flows from operations. TD is total diversification. AGE is the natural logarithm of firm age. SGROWTH is the average sales growth over the prior three fiscal years. LEV is the ratio of total debt to shareholders' equity. EPS is earnings per share. ΔEPS is the year-to-year change in EPS. GENDIFF is the signed difference between the observed ratio of female board members in a given firm and the expected ratio of female board representation. ABSGENDIFF is the absolute value of GENDIFF. NEGGAP is an indicator variable equal to 1 if GENDIFF is negativ, 0 otherwise. See Appendix A for detailed variable descriptions.

4.3 Correlations

We present Pearson correlations in Table 35. Consistent with our descriptive statistics, female board presence is positively associated with RET, ROA, SIZE, BOUTSIDE, TD and LEV, but negatively associated with RISK, AGE and SGROWTH (Table 35). Moreover, we calculate Variance Inflations Factors (VIFs) to control for multicollinearity in a pooled cross-section. Additionally, we estimate VIFs for all variables (Model 1) as well as for each model (Equation 8 – 11, Model 2 – 5), separately (Table 36). All VIFs are far below conventional levels.

Table 35
Pearson Correlations

	Variable	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14	(15)	(16)
(1)	RET	1.00															
(2)	GENDER	**0.05**	1.00														
(3)	GENDERMATCH	-0.02	1.00	1.00													
(4)	ROA	**0.18**	**0.09**	-0.01	1.00												
(5)	SIZE	**-0.05**	**0.17**	-0.03	**-0.17**	1.00											
(6)	BOUTSIDE	**-0.04**	**0.14**	-0.04	**-0.04**	**0.46**	1.00										
(7)	RISK	0.02	**-0.05**	0.01	**0.07**	**-0.26**	**-0.16**	1.00									
(8)	TD	**0.02**	**0.03**	**-0.04**	**-0.03**	**0.16**	**0.12**	**-0.10**	1.00								
(9)	AGE	**-0.03**	**-0.11**	-0.01	**-0.03**	**0.13**	**0.21**	**-0.13**	**0.10**	1.00							
(10)	SGROWTH	**-0.07**	**-0.01**	0.01	0.01	**-0.05**	**-0.06**	**0.34**	**-0.05**	**-0.13**	1.00						
(11)	LEV	**-0.03**	**0.08**	0.01	**-0.21**	**0.48**	**0.16**	**-0.08**	**0.02**	0.01	-0.01	1.00					
(12)	EPS	**0.36**	**0.07**	**-0.07**	**0.59**	**0.11**	**0.06**	**-0.08**	**0.04**	**0.02**	**-0.06**	-0.01	1.00				
(13)	ΔEPS	**0.05**	0.02	0.00	**0.04**	-0.01	**0.02**	-0.01	-0.00	0.01	**-0.03**	-0.01	**0.07**	1.00			
(14)	ABSGENDIFF	**0.04**	**0.13**	**-0.26**	**0.08**	**-0.09**	**-0.16**	**0.04**	**-0.05**	**-0.17**	**0.05**	0.01	**0.03**	0.01	1.00		
(15)	NEGGAP	-0.01	**-0.55**	**-0.74**	**-0.08**	**-0.14**	**-0.07**	0.01	**-0.04**	0.01	0.01	**-0.06**	**-0.03**	-0.02	**-0.21**	1.00	
(16)	ABSGENDIFF × NEGGAP	**0.04**	**-0.33**	**-0.70**	0.02	**-0.20**	**-0.21**	**0.08**	**-0.08**	**-0.20**	**0.11**	**-0.05**	0.01	-0.01	**0.47**	**0.52**	1.00

Table 35 presents pairwise Pearson correlations for our full sample (n = 8,872). GENDER is an indicator variable equal to 1 if a firm is included into the treatment group (has at least one female board member), 0 otherwise. GENDERMATCH is an indicator variable equal to 1 if a firm is included into the treatment group after applying the propensity score matching, 0 otherwise. RET is stock returns. ROA is return on assets. SIZE is the natural logarithm of total assets. BOUTSIDE is the natural logarithm of outside directors. RISK is the standard deviation of cash flows from operations. TD is total diversification. AGE is the natural logarithm of firm age. SGROWTH is the average sales growth over the prior three fiscal years. LEV is the ratio of total debt to shareholders' equity. EPS is earnings per share. ΔEPS is the year-to-year change in EPS. ABSGENDIFF is the absolute value of GENDIFF. NEGGAP is an indicator variable equal to 1 if GENDIFF is negativ, 0 otherwise. GENDIFF is the signed difference between the observed ratio of female board members in a given firm and the expected ratio of female board representation. See Appendix A for detailed variable descriptions. Bold font indicates significance at the two-tailed 10 %-level or higher.

Table 36
Variance Inflation Factors

Variables	Model 1 Pooled	Model 2 PSM Model	Model 3 VR-Model Unmatched	Model 4 VR-Model Matched	Model 5 Gender Expect. Model
GENDER	2.90		1.00		
GENDERMATCH	2.90			1.00	
ROA	1.89	1.08			
SIZE	1.75	1.30			
BOUTSIDE	1.40	1.17			
RISK	1.17	1.19			
TD	1.08	1.03			
AGE	1.22	1.13			
SGROWTH	1.13	1.14			
LEV	1.40	1.11			
EPS	1.82		1.01	1.02	1.01
ΔEPS	1.02		1.00	1.01	1.00
ABSGENDIFF	2.92				2.04
NEGGAP	3.34				2.17
ABSGENDIFF × NEGGAP	5.87				2.66

Table 36 presents variance inflation factors (VIFs) for all variables in the pooled sample (Model 1) as well as separately for each Equation (Equations 8 – 11) of our actual analyses (Model 2 – Model 5). GENDER is an indicator variable equal to 1 if a firm is included into the treatment group (has at least one female board member), 0 otherwise. GENDERMATCH is an indicator variable equal to 1 if a firm is included into the treatment group after applying the PSM model, 0 otherwise. ROA is return on assets. SIZE is the natural logarithm of total assets. BOUTSIDE is the natural logarithm of outside directors. RISK is the standard deviation of cash flows from operations. TD is total diversification. AGE is the natural logarithm of firm age. SGROWTH is the average sales growth over the prior three fiscal years. LEV is the ratio of total debt to shareholders' equity. EPS is earnings per share. ΔEPS is the year-to-year change in EPS. ABSGENDIFF is the absolute value of GENDIFF. NEGGAP is an indicator variable equal to 1 if GENDIFF is negative, 0 otherwise. GENDIFF is the signed difference between the observed ratio of female board members in a given firm and the expected ratio of female board representation. See Appendix A for detailed variable descriptions.

While the VIF for the interaction term in the pooled cross-section is 5.87, which is still far below the critical level of 10, it is much smaller in the model in which it is actually employed (Table 36, Model 5). Therefore, we conclude that multicollinearity does not bias our analyses.

5 Results

5.1 Propensity Score Matching

In the following, we aim to determine whether firms with gender diverse boards exhibit higher long-term stock returns compared to their counterparts. In a first step to test H1, we present univariate t-tests of differences in means for all variables included in our prediction model as well as stock returns (Table 37).

Table 37
Descriptive Statistics – Unmatched Sample

	No female director (GENDER = 0; N = 3,301)			At least one female director (GENDER = 1; N = 5,571)			Difference in means
Variables	Mean	Median	SD	Mean	Median	SD	p-value
RET	0.089	0.003	0.514	0.137	0.084	0.509	<0.001
ROA	0.033	0.028	0.057	0.044	0.040	0.063	<0.001
SIZE	15.79	15.61	1.36	16.35	16.14	1.667	<0.001
BOUTSIDE (ln)	1.743	1.792	0.586	1.896	1.946	0.476	<0.001
BOUTSIDE (#)	6	6	2	7	7	2	
RISK	0.038	0.028	0.037	0.034	0.025	0.033	<0.001
TD	0.859	0.831	0.490	0.888	0.891	0.447	0.006
AGE (ln)	3.809	4.094	0.813	3.617	3.664	0.845	<0.001
AGE (#)	45	60	2	37	39	2	
SGROWTH	0.082	0.042	0.257	0.077	0.054	0.163	0.327
LEV	2.596	1.248	4.509	3.589	1.694	6.191	<0.001

Table 37 presents summary statistics and comparisons in means between firms with at least one female director (GENDER = 1) and without female directors (GENDER = 0). Presented p-values are based on two tailed t-tests for differences in means. GENDER is an indicator variable equal to 1 if a firm is included into the treatment group (has at least one female board member), 0 otherwise. RET is stock returns. ROA is return on assets. SIZE is the natural logarithm of total assets. BOUTSIDE is the natural logarithm of outside directors. RISK is the standard deviation of cash flows from operations. TD is total diversification. AGE is the natural logarithm of firm age. SGROWTH is the average sales growth over the prior three fiscal years. LEV is the ratio of total debt to shareholders' equity. See Appendix A for detailed variable descriptions.

Consistent with our descriptive statistics, we find that before the matching process most differences are highly significant. Thus, we conclude that firms with gender diverse boards are systematically different from firms with all-male boards. Moreover, it seems that firms with female board presence exhibit significantly higher stock returns.

Nevertheless, to make a more convinced statement about the effect of gender diversity, we require a treatment and control group with similar firm fundamentals. To correct for the apparent self-selection, we employ PSM. Afterwards, firms of both groups should be similar with respect to firm characteristics that determine female board representation. Remaining differences in stock returns could, therefore, be attributed to differences in gender diversity.

We present results of our prediction/matching model (Equation 8) for each country in Appendix B. These indicate that the matching process is highly effective. First, the post-match pseudo-R^2 values are significantly smaller than their pre-match counterparts. Second, post-match almost none of the explanatory variables still exhibits a significant association with the treatment. Hence, the matching process balances differences across covariates. Across the different pre-match country subsamples especially SIZE, AGE, SGROWTH, and BOUTSIDE largely have significant predictive ability towards receiving treatment. Overall, coefficients' signs and estimated magnitudes are consistent with prior literature (Gul et al. 2011; Hillman et al. 2007). As such, the prediction model, although developed for a U.S. setting, performs well internationally.

Table 38 presents descriptive statistics for the matched sample as well as results of two-sided t-tests for post-match differences in means. In comparison to the descriptive statistics presented in Table 34 and Table 37, it again appears that the matching process is highly efficient. With the exception of TD and BOUTSIDE, no significant differences remain. Hence, we have first indication that H1 cannot be confirmed. Properly controlling for endogeneity and self-selection seems to render stock market performance differences insignificant.

Table 38
Descriptive Statistics – Matched Sample

	No female director (GENDERMATCH = 0; N = 1,129)			At least one female director (GENDERMATCH = 1; N = 1,129)			Difference in means
Variables	Mean	Median	SD	Mean	Median	SD	p-value
RET	0.137	0.064	0.552	0.116	0.041	0.549	0.371
ROA	0.043	0.036	0.065	0.042	0.041	0.072	0.745
SIZE	15.58	15.46	1.383	15.49	15.33	1.479	0.501
BOUTSIDE	1.662	1.609	0.570	1.620	1.609	0.543	0.073
RISK	0.043	0.031	0.040	0.044	0.032	0.039	0.594
TD	0.878	0.849	0.466	0.839	0.825	0.470	0.049
AGE	3.536	3.584	0.866	3.515	3.611	0.897	0.580
SGROWTH	0.110	0.078	0.209	0.110	0.069	0.233	0.967
LEV	2.628	1.193	5.260	2.783	1.334	5.518	0.495

Table 38 presents summary statistics and comparisons in means for the matched sample. Presented p-values result from two tailed t-tests for differences in means. GENDERMATCH is an indicator variable equal to 1 if a firm is included into the treatment group after applying the propensity score matching, 0 otherwise. RET is stock returns. ROA is return on assets. SIZE is the natural logarithm of total assets. BOUTSIDE is the natural logarithm of outside directors. RISK is the standard deviation of cash flows from operations. TD is total diversification. AGE is the natural logarithm of firm age. SGROWTH is the average sales growth over the prior three fiscal years. LEV is the ratio of total debt to shareholders' equity. See Appendix A for detailed variable descriptions.

As the average pre-match pseudo-R^2 amounts to about 33 percentage (Appendix B), other factors likely influence the propensity of board gender diversity. If the matches were (nearly) perfect, simple univariate t-tests for our variable of interest would suffice to disprove the hypothesis (Dehejia and Wahba 2002; Heckman et al. 1997). However, in the following, we employ a multivariate regression design to control for any remaining differences.

5.2 Value Relevance Design

Table 39 provides the regression results for our value relevance design (Equation 9 and 10). Both investigate firms' stock returns, which are explained by earnings per share (EPS) and the change in earnings per share (ΔEPS). Moreover, we conduct our analyses for the unmatched (GENDER) and matched (GENDERMATCH) samples to exhibit the incremental impact of board gender diversity on stock returns.

Table 39
Value Relevance Regression

	Model 1	Model 2	Model 3	Model 4
	Unmatched Sample	Matched Sample	Unmatched Sample	Matched Sample
GENDER	0.024 **		0.027 ***	
	(3.04)		(3.48)	
GENDERMATCH		0.005		0.004
		(0.30)		(0.23)
EPS	1.221 ***	1.261 ***	1.333 ***	1.193 **
	(10.18)	(4.64)	(8.47)	(2.88)
ΔEPS	0.023	-0.004	0.022	-0.003
	(1.42)	(-0.22)	(1.37)	(-0.18)
ROA			-0.185	0.196
			(-1.32)	(0.55)
SIZE			-0.021 ***	-0.015
			(-5.98)	(-1.57)
BOUTSIDE			0.009	0.034
			(0.96)	(1.59)
RISK			0.449 *	0.544
			(2.29)	(1.22)
TD			0.019 *	0.001
			(2.54)	(0.02)
AGE			-0.022 ***	-0.038 ***
			(-4.86)	(-3.79)
SGROWTH			-0.042	-0.079
			(-1.68)	(-1.54)
LEV			0.001	-0.001
			(1.24)	(-0.14)
Constant	0.097 ***	0.066	0.480 ***	0.384 **
	(5.56)	(1.36)	(8.17)	(2.62)
Fixed Effects	Y,I	Y,I	Y,I	Y,I
R^2	0.453	0.433	0.458	0.439
N	8,872	2,258	8,872	2,258

Table 39 presents results of OLS Value Relevance regressions (Equations 9 and 10) for the unmatched and matched sample (Model 1 and 2). In addition, all control variables from Equation 8 are included in Model 3 and 4 to control for potentially remaining post-match differences. Dependent variable is RET. RET is stock returns. GENDER is an indicator variable equal to 1 if a firm is included into the treatment group (has at least one female board member), 0 otherwise. GENDERMATCH is an indicator variable equal to 1 if a firm is included into the treatment group after applying the PSM model, 0 otherwise. EPS is earnings per share. ΔEPS is the year-to-year change in EPS. ROA is return on assets. SIZE is the natural logarithm of total assets. BOUTSIDE is the natural logarithm of outside directors. RISK is the standard deviation of cash flows from operations. TD is total diversification. AGE is the natural logarithm of firm age. SGROWTH is the average sales growth over the prior three fiscal years. LEV is the ratio of total debt to shareholders' equity. See Appendix A for detailed variable descriptions. Standard errors are clustered on the firm-level (two-tailed t-statistics in parentheses). Each model is estimated with year- (Y) and industry- (I) fixed effects. ***, **, and * denote significance at the 1 %, 5 %, and 10 % level, respectively.

As can be seen in the first column, the coefficient on the GENDER indicator is significantly positive at the 5 %-level ($p < 0.05$). This confirms the pre-match univariate results (Table 37). In the second column, we present results focusing on matched pairs. The coefficient on GENDERMATCH is still positive but highly insignificant. This implies that, by only considering comparable firms, female board representation has no incremental predictive value for stock returns. To alleviate concerns that PSM might not have sufficiently addressed differences between the treatment and control group, we follow Dehejia and Wahba (2002) and include the full set of matching covariates as additional controls (Model 3 and 4). Nevertheless, all inferences remain unchanged.

Summing up, the results from comparing all board gender diverse firms with their counterparts reveal that GENDER has a significantly positive association with stock returns. Considering only matched pairs (i.e., a more comparable set of firms), the significant association vanishes. Hence, long-term changes in shareholder wealth seem independent from board gender diversity *per se*. To put it differently, just appointing (more) women to corporate boards (for instance, due to demand-side regulation), without putting in place additional good governance mechanisms and procedures, will likely be insufficient from the perspective of shareholders.

5.3 Gender Diversity Expectation Model

Table 40 reports results of our multivariate cross-sectional gender diversity expectation model employing the unmatched sample. In the first column, we present the pooled sample (Model 1). Furthermore, we conduct several additional tests. First, we split our sample according to the median of overall societal gender equality, based on World Economic Forum's Global Gender Gap Index (GGGI; Model 2). Second, we split the sample into firms which explicitly formulated a board gender diversity policy (BOARD_POL), and those that did not (Model 3). We follow Rhode and Packel (2014) and argue that corporations with a commitment to gender diversity have access to a broader talent pool with diverse leadership skills. To identify firms with gender diverse policies, we use the variable CGBSDP0013 from the Asset4 database. It presents an indicator whether a firm has committed itself to a policy of board gender diversity. Third, we split the sample into industries in which women are historically more highly represented on corporate boards (i.e., service-oriented, labor-intensive and consumer product industries (Cumming et al. 2015; Farrell and Hersch 2005; Harrigan 1981) and their counterparts (Model 4).

Table 40

Gender Diversity Expectation Model – Unmatched Sample

	Model 1	Model 2		Model 3		Model 4	
	Pooled Sample	GGGI High	GGGI Low	BOARD_POL Y	BOARD_POL N	Gender Industry	Non-Gender Industry
ABSGENDIFF	0.001	-0.001	0.001	-0.001	0.001	-0.001	0.001
	(-0.45)	(-0.53)	(-0.31)	(-0.97)	(0.12)	(-0.66)	(0.08)
NEGGAP	-0.030 ***	0.006	-0.046 ***	-0.014	-0.033 ***	-0.043 ***	-0.017
	(-2.84)	-0.31	(-3.05)	(-0.64)	(-2.75)	(-2.79)	(-1.21)
ABSGENDIFF × NEGGAP	0.005 ***	0.002	0.008 ***	0.004	0.006 ***	0.006 ***	0.004 **
	(4.01)	(0.99)	(3.29)	(1.64)	(3.65)	(3.46)	(2.25)
EPS	1.224 ***	1.613 ***	0.864 ***	1.403 ***	1.167 ***	1.136 ***	1.281 ***
	(10.26)	(11.11)	(5.57)	(7.23)	(8.20)	(5.99)	(8.34)
ΔEPS	0.023	0.011	0.028	0.020	0.023	0.026	0.021
	(1.43)	(0.66)	(1.27)	(0.85)	(1.23)	(1.15)	(0.95)
Constant	0.001	-0.001	0.001	-0.001	0.001	-0.001	0.001
	(-0.45)	(-0.53)	(-0.31)	(-0.97)	(0.12)	(-0.66)	-0.08
Fixed Effects	Y, I	Y, I	Y, I	Y, I	Y, I	Y, I	Y, I
R^2	0.453	0.438	0.475	0.391	0.475	0.435	0.471
N	8,872	4,460	4,412	2,224	6,648	4,114	4,758

Table 40 presents results of the Gender Diversity Expectation Model (Equation 11). We estimate the incremental effect on stock returns (RET), if the actual ratio of female board members downwardly deviates from the average country-industry-level ratio employing the *unmatched* sample (GENDER 1 and 0). Model 1 presents regression results for the pooled sample. Model 2 presents results by splitting the sample according to the median of societal gender equality (GGGI). Model 3 presents results by splitting the sample into firms with (Y) and without (N) board gender diversity policies (BOARD_POL). Model 4 presents results by splitting the sample into industries in which women are historically (under)represented. Dependent variable is RET. RET is stock returns. EPS is earnings per share. ΔEPS is the year-to-year change in EPS. ABSGENDIFF is the absolute value of GENDIFF. NEGGAP is an indicator variable equal to 1 if GENDIFF is negative, 0 otherwise. GENDIFF is the signed difference between the observed ratio of female board members in a given firm and the expected ratio of female board representation. See Appendix A for detailed variable descriptions. Standard errors are clustered on the firm-level (two-tailed t-statistics in parentheses). Each model is estimated with year- (Y) and industry- (I) fixed effects. ***, **, and * denote significance at the 1 %, 5 %, and 10 % level, respectively.

Table 41

Gender Diversity Expectation Model – Matched Sample

	Model 1	Model 2		Model 3		Model 4	
	Pooled Sample	GGGI High	GGGI Low	BOARD_POL Y	BOARD_POL N	Gender Industry	Non-Gender Industry
ABSGENDIFF	-0.003	-0.005	-0.001	-0.004	-0.001	-0.006	-0.001
	(-1.14)	(-1.48)	(-0.24)	(-0.75)	(-0.53)	(-1.26)	(-0.53)
NEGGAP	-0.020	-0.001	-0.028	-0.012	-0.01	-0.120 **	0.034
	(-0.68)	(-0.01)	(-0.67)	(-0.15)	(-0.33)	(-1.99)	-1.05
ABSGENDIFF × NEGGAP	0.004	0.003	0.004	0.002	0.004	0.011 **	-0.001 **
	(1.40)	(0.62)	(0.88)	(0.26)	(1.28)	(2.07)	(-0.09)
EPS	1.257 ***	1.611 ***	0.875 **	1.826 ***	1.073 ***	0.506 ***	1.717 ***
	(4.60)	(5.82)	(2.07)	(5.34)	(3.27)	(1.04)	(9.10)
ΔEPS	-0.004	0.003	-0.009	0.006	-0.007	-0.012	-0.002
	(-0.19)	(0.14)	(-0.28)	(0.13)	(-0.32)	(-0.17)	(-0.17)
Constant	-0.003	-0.005	-0.001	-0.004	-0.001	-0.006	-0.001
	(-1.14)	(-1.48)	(-0.24)	(-0.75)	(-0.53)	(-1.26)	(-0.53)
Fixed Effects	Y, I	Y, I	Y, I	Y, I	Y, I	Y, I	Y, I
R^2	0.433	0.460	0.431	0.362	0.466	0.375	0.497
N	2,258	1,150	1,108	479	1,779	808	1,450

Table 41 presents results of the Gender Diversity Expectation Model (Equation 11). We estimate the incremental effect on stock returns (RET), if the actual ratio of female board members downwardly deviates from the average country-industry-level ratio employing the *matched* sample (GENDERMATCH 1 and 0). Model 1 presents regression results for the pooled sample. Model 2 presents results by splitting the sample according to the median of societal gender equality (GGGI). Model 3 presents results by splitting the sample into firms with (Y) and without (N) board gender diversity policies (BOARD_POL). Model 4 presents results by splitting the sample into industries in which women are historically (under)represented. Dependent variable is RET. RET is stock returns. EPS is earnings per share. ΔEPS is the year-to-year change in EPS. ABSGENDIFF is the absolute value of GENDIFF. NEGGAP is an indicator variable equal to 1 if GENDIFF is negative, 0 otherwise. GENDIFF is the signed difference between the observed ratio of female board members in a given firm and the expected ratio of female board representation. See Appendix A for detailed variable descriptions. Standard errors are clustered on the firm-level (two-tailed t-statistics in parentheses). Each model is estimated with year- (Y) and industry- (I) fixed effects. ***, **, and * denote significance at the 1 %, 5 %, and 10 % level, respectively.

The coefficient on the interaction term ABSGENDIFF × NEGGAP is significantly positive in the pooled sample (Table 40, Model 1). This indicates that the greater the actual ratio of female board members is below expectations, the more positive is the incremental effect on stock returns. Yet, the result seems primarily driven by firms located in countries with lower societal gender equality (Table 40, Model 2). On the surface, this might be in line with role congruity theory. In countries, which generally put less emphasis on empowering women, investors seem to punish firms for appointing female directors. Nevertheless, again attributable to endogeneity concerns, the unmatched sample may result in incorrect inferences. Therefore, Table 41 reports results based on the matched sample. For the pooled sample, the coefficient of the interaction term is now profoundly insignificant. In line with our previous results, board gender diversity exhibits no incremental (negative) effect on stock returns by only considering matched firms.

Referring to the generally insignificant coefficient of the interaction term (ABSGENDIFF × NEGGAP) across the cross-sectional sub-samples, we are able to confirm that firms do not exhibit impaired (nor improved) equity capital market performance if they downwardly deviate from the expected ratio of board gender diversity. While prior research finds that a gender-friendly societal climate seems to translate, on average, into increased female board participation (e.g., Loy and Rupertus 2018a), our results indicate that firms which do not conform to these expectations do not seem to get significantly punished (or rewarded) by investors.

5.4 Robustness Checks

We conduct several sensitivity tests. Our first set of robustness checks (I.I through I.III) refers to the definition of GENDER in our main results. There, we define GENDER as an indicator variable equal to 1 if a firm has at least one female board member, and 0 otherwise. However, using this indicator variable does not allow to make a convinced statement about the effect of adding more women. Prior research suggests that an unimpressive number of female board members (for instance, only one) would not receive great attention and only serve as a 'token' (Adams and Ferreira 2009; Rhode and Packel 2014). Hence, we include three additional robustness checks. First, we compare firms with a very high number of female board members (i.e., the top quartile of the distribution of GENQUOT) with firms with a low degree of board gender diversity (i.e., the bottom quartile). All our main inferences remain unchanged (Table 42 and Table

43, Panel A). Second, we match firms with a female board ratio of at least 25 percent with firms that have no female board members. Again, all inferences from our main analysis hold (Table 42 and Table 43, Panel B). Third, instead of binary indicator variables, we include the observed ratio of female board members as our main independent variable in the value relevance regression and re-run Equation 9 and 10. Both regressions yield similar results to those presented in our main analysis (Table 43, Panel C). Thus, we conclude that besides the effect regarding the presence of female board members, adding more women does not seem to have an impact on a firm's stock performance.

Table 42
Robustness Check I – Alternative Measurement of GENDER

	Panel A Robustness Check I.I High GQ ≥ p(75) & Low GQ ≤ p(25)			Panel B Robustness Check I.II High GQ ≥ 25 % & Low GQ = 0 %		
	GENDER-MATCH = 0	GENDER-MATCH = 1	p-value	GENDER-MATCH = 0	GENDER-MATCH = 1	p-value
RET	0.137	0.064	0.552	0.116	0.041	0.549
ROA	0.043	0.036	0.065	0.042	0.041	0.072
SIZE	15.58	15.46	1.383	15.49	15.33	1.479
BOUTSIDE	1.662	1.609	0.570	1.620	1.609	0.543
RISK	0.043	0.031	0.040	0.044	0.032	0.039
TD	0.878	0.849	0.466	0.839	0.825	0.470
AGE	3.536	3.584	0.866	3.515	3.611	0.897
SGROWTH	0.110	0.078	0.209	0.110	0.069	0.233
LEV	2.628	1.193	5.260	2.783	1.334	5.518
EPS	0.054	0.057	0.094	0.041	0.053	0.102
ΔEPS	-0.021	-0.001	0.891	-0.027	-0.002	0.937

Table 42 presents comparison in means using three alternative measures for board gender diversity. Panel A presents results by matching firms with board gender diversity in the top quartile with firms with board gender diversity in the bottom quartile of the distribution. Panel B presents results by matching firms with female board representation of more than 25 percent with firms without any women on their boards. GQ is the observed ratio of female board members (GENQUOT). GENDERMATCH is an indicator variable equal to 1 if a firm is included into the treatment group after applying the PSM model, 0 otherwise. RET is stock returns. ROA is return on assets. SIZE is the natural logarithm of total assets. BOUTSIDE is the natural logarithm of outside directors. RISK is the standard deviation of cash flows from operations. TD is total diversification. AGE is the natural logarithm of firm age. SGROWTH is the average sales growth over the prior three fiscal years. LEV is the ratio of total debt to shareholders' equity. EPS is earnings per share. ΔEPS is the year-to-year change in EPS. Presented p-values are based on two tailed t-tests for differences in means. See Appendix A for detailed variable descriptions.

Table 43

Robustness Check I – Alternative Measurements of GENDER – Multivariate Results

	Panel A Robustness Check I.I High GQ ≥ p(75) & Low GQ ≤ p(25)				Panel B Robustness Check I.II High GQ ≥ 25 % & Low GQ = 0 %				Panel C Robustness Check I.III GENQUOT	
	Model 1		Model 2		Model 1		Model 2		Model 1	
	Unmatched Sample	Matched Sample	Unmatched Sample	Matched Sample	Unmatched Sample	Matched Sample	Unmatched Sample	Matched Sample	Unmatched Sample	Matched Sample
GENDER	0.002				0.007					
	(0.19)				(0.55)					
GENDERMATCH		0.014				0.020				
		(0.95)				(0.58)				
GENQUOT									0.001 **	0.001
									(2.47)	(0.53)
ABSGENDIFF			0.001	0.001			0.000	-0.008		
			(-0.44)	(-0.05)			(-0.32)	(-1.63)		
NEGGAP			-0.038 ***	-0.034			0.173 ***	-0.059		
			(-2.74)	(-1.24)			(4.33)	(-0.46)		
ABSGENDIFF × NEGGAP			0.006 ***	0.003			-0.009 ***	0.002		
			(4.35)	(1.12)			(-2.80)	(0.26)		
EPS	1.256 ***	1.562 ***	1.249 ***	1.561 ***	1.327 ***	1.691 ***	1.304 ***	1.734 ***	1.223 ***	1.261 ***
	(8.57)	(7.43)	(8.58)	(7.43)	(7.23)	(3.15)	(7.22)	(3.26)	(10.21)	(4.64)
ΔEPS	0.031	0.006	0.032	0.006	0.035	0.007	0.033	0.005	0.023	-0.004
	(1.10)	(0.32)	(1.12)	(0.32)	(1.03)	(0.23)	(0.97)	(0.18)	(1.42)	(-0.23)

Table 43

Robustness Check I – Alternative Measurements of GENDER – Multivariate Results – continued

	Panel A Robustness Check I.I High GQ ≥ p(75) & Low GQ ≤ p(25)				Panel B Robustness Check I.II High GQ ≥ 25 % & Low GQ = 0 %				Panel C Robustness Check I.III GENQUOT	
	Model 1		Model 2		Model 1		Model 2		Model 1	
	Unmatched Sample	Matched Sample	Unmatched Sample	Matched Sample	Unmatched Sample	Matched Sample	Unmatched Sample	Matched Sample	Unmatched Sample	Matched Sample
Constant	0.106 ***	-0.03	0.115 ***	-0.018	0.087 ***	-0.151 *	0.140 ***	-0.020	0.102 ***	0.064
	(5.03)	(-0.79)	(4.94)	(-0.45)	(3.71)	(-1.75)	(3.12)	(-0.17)	(5.86)	(1.33)
Fixed Effects	Y, I	Y, I	Y,I	Y,I	Y, I	Y, I	Y,I	Y,I	Y, I	Y, I
R^2	0.434	0.478	0.436	0.478	0.421	0.433	0.428	0.435	0.452	0.433
N	5,881	2,076	5,881	2,076	3,907	552	3,907	552	8,872	2,258

Table 43 presents results of OLS Value Relevance regressions (Model 1, Equation 9 and 10) as well as the Gender Diversity Expectation Model (Model 2, Equation 11) for the matched and unmatched sample. Panel A presents results by matching firms with board gender diversity in the top quartile with firms with board gender diversity in the bottom quartile of the distribution. Panel B presents results by matching firms with female board representation of more than 25 percent with firms without any women on their boards. Panel C presents results including GENQUOT instead of GENDER/GENDERMATCH. GQ is the observed ratio of female board members (GENQUOT). Dependent variable is RET. Ret is stock returns. GENDER is an indicator variable equal to 1 if a firm is included into the treatment group (has at least one female board member), 0 otherwise. GENDERMATCH is an indicator variable equal to 1 if a firm is included into the treatment group after applying the PSM model, 0 otherwise. ABSGENDIFF is the absolute value of GENDIFF. NEGGAP is an indicator variable equal to 1 if GENDIFF is negative, 0 otherwise. GENDIFF is the signed difference between the observed ratio of female board members in a given firm and the expected ratio of female board representation. EPS is earnings per share. ΔEPS is the year-to-year change in EPS. See Appendix A for detailed variable descriptions. Standard errors are clustered at the firm-level (two-tailed t-statistics in parentheses). ***, **, and * denote significance at the 1 %, 5 %, and 10 % level, respectively.

Our second set of robustness checks (II.I through II.III) adjusts our dependent variable (i.e., stock returns). First, as we are also interested in the long-term effect of gender diverse boards, we additionally employ three-year buy-and-hold returns (BHR). This builds on the notion that some new strategies take more time to implement and come into effect (e.g., Pan et al. 2016). While one could also consider even longer periods, at significant data loss, three years are probably sufficient for market participants to finally assess board members' quality. The inferences from our main analysis remain virtually unchanged (Table 44 and Panel A of Table 45). Moreover, we use market- (MAR) as well as industry-adjusted stock returns (IAR) to provide two alternative measures of changes in shareholder value. On the one hand, we base our market return on the return of the top index of each country (Table 45, Panel B). On the other hand, industry returns are estimated as the average stock return in each industry per year. Hence, we subtract either the market or the industry return from a firm's raw stock return and re-perform our analysis. Results from the two-sided t-tests (untabulated), value relevance regression our analysis as well as from the Gender Diversity Expectation Model remain qualitatively unchanged (Table 45, Panel C).

Table 44

Robustness Check II – Alternative Measurements of RET – Univariate Results

	Robustness Check II.I Buy-and-Hold Returns (BHR)		
	GENDER-MATCH = 0	GENDER-MATCH = 1	
	Mean	Mean	p-value
BHR	0.269	0.213	0.182
ROA	0.046	0.047	0.662
SIZE	16.66	16.54	0.440
BOUTSIDE	1.681	1.649	0.266
RISK	0.042	0.043	0.643
TD	0.870	0.854	0.553
AGE	3.493	3.466	0.589
SGROWTH	0.106	0.106	0.970
LEV	2.623	2.244	0.201
N	606	606	

Table 44 presents comparisons in means for the buy-and-hold return (BHR) over three years as well as three-year averages of the other variables. Presented p-values are based on two tailed t-tests for differences in means. GENDERMATCH is an indicator variable equal to 1 if a firm is included into the treatment group after applying the PSM model, 0 otherwise. ROA is return on assets. SIZE is the natural logarithm of total assets. BOUTSIDE is the natural logarithm of outside directors. RISK is the standard deviation of cash flows from operations. TD is total diversification. AGE is the natural logarithm of firm age. SGROWTH is the average sales growth over the prior three fiscal years. LEV is the ratio of total debt to shareholders' equity. See Appendix A for detailed variable descriptions.

Table 45

Robustness Check II – Alternative Measurements of RET – Multivariate Results

	Panel A Robustness Check II.I Buy-and-Hold Returns (BHR)				Panel B Robustness Check II.II Market-Adjusted Returns (MAR)			
	Model 1		Model 2		Model 1		Model 2	
	Unmatched Sample	Matched Sample	Unmatched Sample	Matched Sample	Unmatched Sample	Matched Sample	Unmatched Sample	Matched Sample
GENDER	0.058 **				0.022 ***			
	(2.03)				(2.84)			
GENDERMATCH		-0.024				0.000		
		(-0.54)				(0.02)		
ABSGENDIFF			-0.001	-0.004			-0.001	-0.003
			(-0.39)	(-0.65)			(-0.73)	(-1.28)
NEGGAP			-0.083 **	-0.035			-0.032 ***	-0.019
			(-2.53)	(-0.49)			(-3.05)	(-0.64)
ABSGENDIFF × NEGGAP			0.016 ***	0.012*			0.006 ***	0.004
			(3.92)	(1.82)			(4.16)	(1.43)
EPS	1.924 ***	1.803 ***	1.916 ***	1.782 ***	1.177 ***	1.206 ***	1.179 ***	1.204 ***
	(10.03)	(5.16)	(10.03)	(5.11)	(9.91)	(4.47)	(9.99)	(4.48)
ΔEPS	0.007	0.028	0.007	0.027	0.022	-0.004	0.022	-0.004
	(0.17)	(0.62)	(0.18)	(0.60)	(1.41)	(-0.22)	(1.42)	(-0.19)
Constant	0.344 ***	-0.037	0.402 ***	-0.045	-0.253 ***	-0.272 ***	-0.229 ***	-0.258 ***
	(6.17)	(-0.32)	(7.35)	(-0.37)	(-13.67)	(-5.44)	(-11.35)	(-4.62)
Fixed Effects	Y, I	Y, I	Y, I	Y, I	Y, I	Y, I	Y, I	Y, I
R^2	0.188	0.202	0.191	0.205	0.536	0.528	0.537	0.528
N	5,230	1,212	5,230	1,212	8,872	2,258	8,872	2,258

Table 45
Robustness Check II – Alternative Measurements of RET – Multivariate Results – continued

	Panel C Robustness Check II.III Industry-Adjusted Returns (IAR)			
	Model 1		Model 2	
	Unmatched Sample	Matched Sample	Unmatched Sample	Matched Sample
GENDER	0.024 **			
	(3.07)			
GENDERMATCH		0.004		
		(0.26)		
ABSGENDIFF			0.001	0.003
			(0.22)	(1.20)
NEGGAP			0.025	0.018
			(2.34)	(0.60)
ABSGENDIFF × NEGGAP			0.005 ***	0.004
			(3.64)	(1.50)
EPS	1.179 ***	1.213 ***	1.182 ***	1.209 ***
	(9.90)	(4.48)	(9.99)	(4.50)
ΔEPS	0.022	-0.003	0.022	-0.002
	(1.37)	(-0.15)	(1.38)	(-0.12)
Constant	-0.088 ***	-0.136 ***	0.066 ***	-0.112 **
	(5.04)	(-2.74)	(3.44)	(-2.16)
Fixed Effects	Y, I	Y, I	Y, I	Y, I
R^2	0.143	0.114	0.191	0.114
N	8,872	2,258	5,291	2,258

Table 45 presents results of OLS Value Relevance regressions (Model 1, Equation 9 and 10) as well as the Gender Diversity Expectation Model (Model 2, Equation 11) using three year buy-and-hold returns (BHR, Panel A), market-adjusted returns (MAR, Panel B) as well as industry-adjusted returns (IAR, Panel C) as the dependent variable. GENDER is an indicator variable equal to 1 if a firm is included into the treatment group (has at least one female board member), 0 otherwise. GENDERMATCH is an indicator variable equal to 1 if a firm is included into the treatment group after applying the propensity score matching, 0 otherwise. ABSGENDIFF is the absolute value of GENDIFF. NEGGAP is an indicator variable equal to 1 if GENDIFF is negative, 0 otherwise. GENDIFF is the signed difference between the observed ratio of female board members in a given firm and the expected ratio of female board representation. EPS is earnings per share. ΔEPS is the year-to-year change in EPS. See Appendix A for detailed variable descriptions. Each model is estimated with year- (Y) and industry- (I) fixed effects. Standard errors are clustered at the firm-level (two-tailed t-statistics in parentheses). ***, **, and * denote significance at the 1 %, 5 %, and 10 % level, respectively.

Our third set of robustness checks (III) acknowledges the fact, that there is an increasing amount of firms that voluntarily commit to governance policies with respect to gender diversity on their boards. Thus, we include BOARD_POL as an additional covariate in our PSM model. Results regarding our first hypothesis remain unchanged (Table 46, Panel A).

Table 46
Robustness Check III – Including BOARD_POL in PSM Model

Panel A: Two-Sided T-Tests

	GENDER-MATCH = 0	GENDER-MATCH = 1	
	Mean	Mean	p-value
RET	0.156	0.122	0.159
ROA	0.043	0.043	0.908
SIZE	16.66	16.56	0.383
BOUTSIDE	1.654	1.607	0.048
RISK	0.043	0.044	0.403
TD	0.864	0.831	0.100
AGE	3.506	3.513	0.845
SGROWTH	0.114	0.111	0.791
LEV	2.692	2.508	0.397
BOARD_POL	0.206	0.191	0.391
N	1,094	1,094	

Panel B: Value Relevance and Gender Expectation Model

	Model 1		Model 2	
	Unmatched Sample	Matched Sample	Unmatched Sample	Matched Sample
GENDER	0.024 **			
	(3.07)			
GENDERMATCH		0.004		
		(0.26)		
ABSGENDIFF			0.001	0.003
			(0.22)	(1.20)
NEGGAP			0.025	0.018
			(2.34)	(0.60)
ABSGENDIFF × NEGGAP			0.005 ***	0.004
			(3.64)	(1.50)
EPS	1.179 ***	1.213 ***	1.182 ***	1.209 ***
	(9.90)	(4.48)	(9.99)	(4.50)
ΔEPS	0.022	-0.003	0.022	-0.002
	(1.37)	(-0.15)	(1.38)	(-0.12)
Constant	-0.088 ***	-0.136 ***	0.066 ***	-0.112 **
	(5.04)	(-2.74)	(3.44)	(-2.16)
Fixed Effects	Y, I	Y, I	Y, I	Y, I
R^2	0.143	0.114	0.191	0.114
N	8,872	2,258	5,291	2,258

Table 46 presents comparisons in means for the matched sample (Panel A) as well as of OLS Value Relevance regressions (Model 1, Equation 9 and 10) and Gender Diversity Expectation Model (Model 2, Equation 11) (Panel B) by adding BOARD_POL to the PSM model. BOARD_POL is an indicator variable equal to 1 if the firm has a policy regarding gender diversity on its board, 0 otherwise. Dependent variable in Panel B is RET. RET is stock returns. Presented p-values in Panel A are based on two tailed t-tests for differences in means. See Appendix A for detailed variable descriptions. Each model in Panel B is estimated with year- (Y) and industry- (I) fixed effects. Standard errors are clustered at the firm-level (two-tailed t-statistics in parentheses). ***, **, and * denote significance at the 1 %, 5 %, and 10 % level, respectively.

Nevertheless, results of our Gender Diversity Expectation Model change slightly. The coefficient on the interaction term ABSGENDIFF × NEGGAP remains significant at a 5 %-level, but is economically rather neglectable (Table 46, Panel B). Moreover, the interaction term is essentially cancelled out by the larger, but insignificant, coefficient on the baseline effect NEGGAP.

In our fourth set of robustness checks (IV), we re-run the PSM procedure only for countries which do not mandate a gender quota by law. We present related results of the two-sided t-tests as well as multivariate analysis in Table 47. However, our main inferences remain unchanged.

Table 47
Robustness Check IV – Only Countries Without a Mandatory Gender Quota

Panel A: Two-Sided T-Tests			
	GENDER-MATCH = 0	GENDER-MATCH = 1	
	Mean	Mean	p-value
RET	0.135	0.115	0.383
ROA	0.043	0.041	0.695
SIZE	16.58	16.50	0.517
BOUTSIDE	1.655	1.608	0.048
RISK	0.043	0.044	0.581
TD	0.876	0.839	0.062
AGE	3.531	3.508	0.535
SGROWTH	0.108	0.108	0.998
LEV	2.608	2.801	0.403
N	1,107	1,107	

Panel B: Value Relevance and Gender Expectation Model				
	Model 1		Model 2	
	Unmatched Sample	Matched Sample	Unmatched Sample	Matched Sample
GENDER	0.024 **			
	(3.03)			
GENDERMATCH		0.006		
		(0.40)		
ABSGENDIFF			0.000	-0.003
			(-0.53)	(-1.34)
NEGGAP			-0.030 ***	-0.026
			(-2.86)	(-0.88)
ABSGENDIFF × NEGGAP			0.005 ***	0.005
			(3.99)	(1.62)

Table 47
Robustness Check IV – Only Countries Without a Mandatory Gender Quota – continued

	Model 1		Model 2	
	Unmatched Sample	Matched Sample	Unmatched Sample	Matched Sample
EPS	1.225 ***	1.250 ***	1.228 ***	1.240 ***
	(10.10)	(4.59)	(10.18)	(4.58)
ΔEPS	0.025	-0.004	0.025	-0.004
	(1.47)	(-0.22)	(1.48)	(-0.19)
Constant	0.128 ***	0.183 ***	0.143 ***	-0.054 **
	(4.92)	(4.37)	(5.23)	(-1.26)
Fixed Effects	Y, I	Y, I	Y,I	Y,I
R^2	0.455	0.114	0.455	0.436
N	8,791	2,258	8,791	2,214

Table 47 presents comparisons in means for the matched sample (Panel A) as well as of OLS Value Relevance regressions (Model 1, Equation 9 and 10) and Gender Diversity Expectation Model (Model 2, Equation 11) (Panel B) with a focus on countries without mandatory gender quotas. Presented p-values in Panel A are based on two tailed t-tests for differences in means. GENDER is an indicator variable equal to 1 if a firm is included into the treatment group (has at least one female board member), 0 otherwise. GENDERMATCH is an indicator variable equal to 1 if a firm is included into the treatment group after applying the PSM model, 0 otherwise. Dependent variable is RET. RET is stock returns. ROA is return on assets. SIZE is the natural logarithm of total assets. BOUTSIDE is the natural logarithm of outside directors. RISK is the standard deviation of cash flows from operations. TD is total diversification. AGE is the natural logarithm of firm age. SGROWTH is the average sales growth over the prior three fiscal years. LEV is the ratio of total debt to shareholders' equity. ABSGENDIFF is the absolute value of GENDIFF. NEGGAP is an indicator variable equal to 1 if GENDIFF is negative, 0 otherwise. GENDIFF is the signed difference between the observed ratio of female board members in a given firm and the expected ratio of female board representation. EPS is earnings per share. ΔEPS is the year-to-year change in EPS. See Appendix A for detailed variable descriptions. Each model in Panel B is estimated with year- (Y) and industry- (I) fixed effects. Standard errors are clustered at the firm-level (two-tailed t-statistics in parentheses). ***, **, and * denote significance at the 1 %, 5 %, and 10 % level, respectively.

Our full sample is heavily skewed towards U.S. and Japanese observations. We address this issue in our fifth set of robustness checks (V.I through V.III). In this context, prior literature suggests that differences in cultural factors and institutional systems might have a significant influence on the variation of female board representation (e.g., Grosvold and Brammer 2011; Grosvold et al. 2016). Thus, we re-run all our regressions once without U.S. observations, once without Japanese observations, and also without observations from both countries (Table 48 and Table 49). As can be seen, all major inferences hold.

Table 48

Robustness Check V – Without the U.S., Japan, or Both – Univariate Results

	Panel A Robustness Check V.I Without the U.S.			Panel B Robustness Check V.II Without Japan			Panel C Robustness Check V.III Without the U.S. & Japan		
	GENDER-MATCH = 0	GENDER-MATCH = 1		GENDER-MATCH = 0	GENDER-MATCH = 1		GENDER-MATCH = 0	GENDER-MATCH = 1	
	Mean	Mean	p-value	Mean	Mean	p-value	Mean	Mean	p-value
RET	0.120	0.090	0.305	0.136	0.114	0.427	0.110	0.073	0.346
ROA	0.038	0.038	0.993	0.046	0.045	0.828	0.040	0.041	0.846
SIZE	17.47	17.49	0.920	15.50	15.37	0.098	15.61	15.58	0.812
BOUTSIDE	1.737	1.704	0.334	1.580	1.544	0.174	1.612	1.596	0.737
RISK	0.041	0.042	0.857	0.046	0.047	0.616	0.048	0.048	0.905
TD	0.935	0.897	0.144	0.834	0.797	0.068	0.879	0.844	0.223
AGE	3.782	3.799	0.711	3.388	3.358	0.473	3.621	3.639	0.761
SGROWTH	0.113	0.113	0.975	0.128	0.127	0.953	0.152	0.149	0.881
LEV	2.871	3.145	0.389	2.711	2.815	0.703	3.412	3.174	0.603
N	681	681		885	885		437	437	

Table 48 presents comparisons in means for the matched sample excluding the U.S. (Panel A), Japan (Panel B) as well as the U.S. and Japan (Panel C) from the final sample. GENDERMATCH is an indicator variable equal to 1 if a firm is included into the treatment group after applying the PSM model, 0 otherwise. ROA is return on assets. SIZE is the natural logarithm of total assets. BOUTSIDE is the natural logarithm of outside directors. RISK is the standard deviation of cash flows from operations. TD is total diversification. AGE is the natural logarithm of firm age. SGROWTH is the average sales growth over the prior three fiscal years. LEV is the ratio of total debt to shareholders' equity. Presented p-values are based on two tailed t-tests for differences in means. See Appendix A for detailed variable descriptions.

Table 49

Robustness Check V – Without the U.S., Japan, or Both – Multivariate Results

	Panel A Robustness Check V.I Without the U.S.				Panel B Robustness Check V.II Without Japan			
	Model 1		Model 2		Model 1		Model 2	
	Unmatched Sample	Matched Sample	Unmatched Sample	Matched Sample	Unmatched Sample	Matched Sample	Unmatched Sample	Matched Sample
GENDER	0.024 ***				-0.014			
	(3.04)				(-1.12)			
GENDERMATCH		-0.008				-0.001		
		(-0.41)				(-0.08)		
ABSGENDIFF			-0.002	-0.002			-0.001	-0.002
			(-1.47)	(-0.89)			(-0.51)	(-0.97)
NEGGAP			-0.032 **	-0.028			0.001	0.001
			(-2.35)	(-0.81)			(0.01)	(0.01)
ABSGENDIFF × NEGGAP			0.008 ***	0.004			0.002	0.002
			(3.68)	(1.03)			(1.51)	(0.69)
EPS	1.221 ***	1.200 ***	1.121 ***	1.206 ***	1.189 ***	1.140 ***	1.190 ***	1.141 ***
	(10.18)	(3.10)	(7.25)	(3.12)	(8.18)	(3.82)	(8.20)	(3.83)
ΔEPS	0.023	-0.003	0.027 *	-0.003	0.024	-0.006	0.024	-0.006
	(1.42)	(-0.16)	(1.68)	(-0.13)	(1.42)	(-0.32)	(1.42)	(-0.30)
Constant	0.097 ***	0.162	0.164 ***	0.169	0.208 ***	-0.040 ***	0.193 ***	-0.037
	(5.56)	(1.53)	(6.09)	(1.49)	(8.27)	(-0.92)	(7.96)	(-0.76)
Fixed Effects	Y, I	Y, I	Y,I	Y,I	Y, I	Y, I	Y,I	Y,I
R^2	0.453	0.414	0.447	0.413	0.472	0.446	0.473	0.446
N	8,872	1,362	5,180	1,362	6,600	1,770	6,600	1,770

Table 49
Robustness Check V – Without the U.S., Japan, or Both – Multivariate Results – continued

	Panel C Robustness Check V.III Without the U.S. & Japan			
	Model 1		Model 2	
	Unmatched Sample	Matched Sample	Unmatched Sample	Matched Sample
GENDER	0.024 **			
	(3.03)			
GENDERMATCH		0.006		
		(0.40)		
ABSGENDIFF			0.000	-0.003
			(-0.53)	(-1.34)
NEGGAP			-0.030 ***	-0.026
			(-2.86)	(-0.88)
ABSGENDIFF × NEGGAP			0.005 ***	0.005
			(3.99)	(1.62)
EPS	1.225 ***	1.250 ***	1.228 ***	1.240 ***
	(10.10)	(4.59)	(10.18)	(4.58)
ΔEPS	0.025	-0.004	0.025	-0.004
	(1.47)	(-0.22)	(1.48)	(-0.19)
Constant	0.128 ***	0.183 ***	0.143 ***	-0.054 **
	(4.92)	(4.37)	(5.23)	(-1.26)
Fixed Effects	Y, I	Y, I	Y,I	Y,I
R^2	0.455	0.114	0.455	0.436
N	8,791	2,258	8,791	2,214

Table 49 presents results of OLS Value Relevance regressions (Model 1, Equation 9 and 10) as well as the Gender Diversity Expectation Model (Model 2, Equation 11) excluding the U.S. (Panel A), Japan (Panel B) as well as the U.S. and Japan (Panel C) from the final sample. Dependent variable is RET. RET is stock returns. GENDER is an indicator variable equal to 1 if a firm is included into the treatment group (has at least one female board member), 0 otherwise. GENDERMATCH is an indicator variable equal to 1 if a firm is included into the treatment group after applying the propensity score matching, 0 otherwise. ABSGENDIFF is the absolute value of GENDIFF. NEGGAP is an indicator variable equal to 1 if GENDIFF is negative, 0 otherwise. GENDIFF is the signed difference between the observed ratio of female board members in a given firm and the expected ratio of female board representation. EPS is earnings per share. ΔEPS is the year-to-year change in EPS. See Appendix A for detailed variable descriptions. Each model in Panel B is estimated with year- (Y) and industry- (I) fixed effects. Standard errors are clustered at the firm-level (two-tailed t-statistics in parentheses). ***, **, and * denote significance at the 1 %, 5 %, and 10 % level, respectively.

Over the course of our period under consideration, several countries in our sample enacted mandatory gender quotas as well as voluntary corporate governance codex guidelines regarding board gender diversity. Thus, in our sixth set of robustness checks (VI), we account for mandatory gender quotas (QUOTA) as well as voluntary corporate governance code stipulations (CORPCODE). Results from our value relevance design as well as the

cross-sectional Gender Diversity Expectation Model are presented in Table 50. Our inferences remain unchanged.

Table 50
Robustness Check VI – QUOTA and CORPCODE

	Model 1		Model 2	
	Unmatched Sample	Matched Sample	Unmatched Sample	Matched Sample
GENDER	0.028 ***			
	(3.57)			
GENDERMATCH		0.005		
		(0.31)		
ABSGENDIFF			-0.001	-0.003
			(-0.64)	(-1.31)
NEGGAP			-0.035 ***	-0.025
			(-3.39)	(-0.86)
ABSGENDIFF × NEGGAP			0.006 ***	0.005
			(4.45)	(1.60)
EPS	1.215 ***	1.256 ***	1.219 ***	1.252 ***
	(10.11)	(4.63)	(10.18)	(4.64)
ΔEPS	0.023	-0.004	0.023	-0.004
	(1.43)	(-0.23)	(1.44)	(-0.20)
QUOTA	-0.113 ***	-0.129 **	-0.112	-0.132 ***
	(-3.52)	(-2.86)	(-3.46)	(-2.91)
CORPCODE	-0.037 **	-0.025	-0.039	-0.031
	(-2.62)	(-0.80)	(-2.79)	(-0.96)
Constant	-0.145 ***	0.307 ***	-0.126 ***	0.331 ***
	(-5.70)	(6.00)	(-4.75)	(5.92)
Fixed Effects	Y, I	Y, I	Y,I	Y,I
R^2	0.455	0.114	0.454	0.436
N	8,791	2,258	8,872	2,214

Table 50 presents results of OLS Value Relevance regression (Model 1, Equation 9 and 10) as well as the Gender Diversity Expectation Model (Model 2, Equation 11) including QUOTA and CORPCODE as additional controls. Dependent variable is RET. RET is stock returns. GENDER is an indicator variable equal to 1 if a firm is included into the treatment group (has at least one female board member), 0 otherwise. GENDERMATCH is an indicator variable equal to 1 if a firm is included into the treatment group after applying the PSM model, 0 otherwise. ABSGENDIFF is the absolute value of GENDIFF. NEGGAP is an indicator variable equal to 1 if GENDIFF is negative, 0 otherwise. GENDIFF is the signed difference between the observed ratio of female board members in a given firm and the expected ratio of female board representation. EPS is earnings per share. ΔEPS is the year-to-year change in EPS. QUOTA is an indicator variable equal to 1 if the country has a mandatory gender quota in year t, 0 otherwise. CORPCODE is an indicator variable equal to 1 if the country has a voluntary corporate governance code stipulation regarding gender, 0 otherwise. See Appendix A for detailed variable descriptions. Each model in Panel B is estimated with year- (Y) and industry- (I) fixed effects. Standard errors are clustered at the firm-level (two-tailed t-statistics in parentheses). ***, **, and * denote significance at the 1 %, 5 %, and 10 % level, respectively.

In our seventh set of robustness checks (VII), to control for any remaining differences between the matched and unmatched sample, we also include all control variables from Equation 8 in our Gender Diversity Expectation Model (Equation 11). Nevertheless, our main results remain unchanged (Table 51).

Table 51
Robustness Check VII – Gender Diversity Expectation Model with Additional Controls

	Gender Diversity Expectation Model	
	Unmatched Sample	Mached Sample
ABSGENDIFF	0.001	-0.002
	(-0.41)	(-0.99)
NEGGAP	-0.027 ***	-0.014
	(-2.63)	(-0.50)
ABSGENDIFF × NEGGAP	0.003 ***	0.003
	(2.60)	(0.86)
EPS	1.329 ***	1.187 ***
	(8.45)	(2.88)
ΔEPS	0.022	-0.003
	(1.38)	(-0.16)
ROA	-0.178	0.206
	(-1.26)	(0.57)
SIZE	-0.020 ***	-0.014
	(-5.51)	(-1.51)
BOUTSIDE	0.014	0.032
	(1.54)	(1.45)
RISK	0.445 **	0.540
	(2.27)	(1.21)
TD	0.019 **	-0.001
	(2.54)	(-0.01)
AGE	-0.022 ***	-0.038 ***
	(-4.97)	(-3.74)
SGROWTH	-0.047 *	-0.079
	(-1.90)	(-1.55)
LEV	0.001	-0.001
	(1.22)	(-0.11)
Constant	0.478 ***	0.396 ***
	(7.92)	(2.72)
Fixed Effects	Y, I	Y, I
R^2	0.457	0.438
N	8,872	2,258

Table 51 presents results for the Gender Diversity Expectation Model (Equation 11) for the unmatched and matched sample, including all control variables from Equation 8. Dependent variable is RET. RET is stock returns. See Appendix A for detailed variable descriptions. Each model is estimated with year- (Y) and industry- (I) fixed effects. Standard errors are clustered at the firm-level (two-tailed t-statistics in parentheses). ***, **, and * denote significance at the 1 %, 5 %, and 10 % level, respectively.

Moreover, we also perform several minor robustness checks. First, the PSM caliper of 0.05 might influence our results. Thus, we reduce it to 0.01. This modification results in a considerably smaller sample, attributable to even less permissible differences in propensity scores, but qualitatively unchanged inferences in all analyses (Table 52 and Table 53, Panel A).

Table 52
Minor Robustness Checks – Univariate Results

	Panel A Robustness Check Caliper 1 %			Panel B Robustness Check Lagged Board Data		
	Matched Sample			Matched Sample		
	GENDER-MATCH = 0	GENDER-MATCH = 1	p-value	GENDER-MATCH = 0	GENDER-MATCH = 1	p-value
RET	0.146	0.142	0.876	0.044	0.042	0.633
ROA	0.044	0.044	0.955	0.044	0.042	0.633
SIZE	16.69	16.68	0.937	16.45	16.35	0.331
BOUTSIDE	1.646	1.633	0.575	1.652	1.626	0.249
RISK	0.043	0.043	0.835	0.042	0.043	0.654
TD	0.865	0.859	0.759	0.851	0.844	0.693
AGE	3.495	3.512	0.663	3.489	3.457	0.383
SGROWTH	0.106	0.101	0.545	0.113	0.114	0.987
LEV	2.524	3.437	0.278	2.628	2.689	0.788
N	981	981		1,201	1,201	

Table 52 presents results from comparisons in means for the matched sample using a caliper of 1 percent (Panel A) and lagged board data (Panel B) in the PSM model. GENDERMATCH is an indicator variable equal to 1 if a firm is included into the treatment group after applying the PSM model, 0 otherwise. RET is stock returns. ROA is return on assets. SIZE is the natural logarithm of total assets. BOUTSIDE is the natural logarithm of outside directors. RISK is the standard deviation of cash flows from operations. TD is total diversification. AGE is the natural logarithm of firm age. SGROWTH is the average sales growth over the prior three fiscal years. LEV is the ratio of total debt to shareholders' equity. See Appendix A for detailed variable descriptions.

Second, we follow prior literature and employ one-year lagged board gender information in our analysis, as the effect of gender diverse boards will most likely occur over time (e.g., Adams and Ferreira 2009; Carter et al. 2010; Pan et al. 2016). Again, this decreases the number of observations but all inferences from the matching procedure for each country, the value relevance regressions, as well as the Gender Diversity Expectation Model remain qualitatively unchanged (Table 52 and Table 53, Panel B). Third, we include country-fixed effects in all our multivariate analysis. As such, we control for country-specific, time-invariant institutional factors. However, our inferences remain robust (Table 53, Panel C).

Table 53

Minor Robustness Checks – Multivariate Results

	Panel A Robustness Check Caliper 1 %				Panel B Robustness Check Lagged Board Data			
	Model 1		Model 2		Model 1		Model 2	
	Unmatched Sample	Matched Sample	Unmatched Sample	Matched Sample	Unmatched Sample	Matched Sample	Unmatched Sample	Matched Sample
GENDER	0.024 ***				0.020 *			
	(3.00)				(2.54)			
GENDERMATCH		0.018				-0.01		
		(1.08)				(-0.41)		
ABSGENDIFF			0.001	-0.001			0.001	0.001
			(-0.49)	(-0.57)			(0.64)	(0.57)
NEGGAP			-0.030 ***	-0.013			-0.030 ***	-0.018
			(-2.80)	(-0.42)			(-2.91)	(-0.66)
ABSGENDIFF × NEGGAP			0.006 ***	0.002			0.005 ***	0.004
			(4.09)	(0.71)			(4.10)	(1.54)
EPS	1.204 ***	1.621 ***	1.208 ***	1.615 ***	1.221 ***	0.957 ***	1.223 ***	0.956 ***
	(9.72)	(8.76)	(9.81)	(8.72)	(10.1)	(3.88)	(10.15)	(3.92)
ΔEPS	0.028	0.012	0.028 *	0.012	0.030	0.040	0.028	0.036
	(1.53)	(0.57)	(0.54)	(0.57)	(1.52)	(1.54)	(1.53)	(1.54)
Constant	0.132 ***	0.189 ***	0.145 ***	-0.169 ***	0.136 ***	-0.030 ***	0.148 ***	-0.035
	(5.00)	(3.91)	(5.21)	(-3.27)	(5.20)	(-0.85)	(5.46)	(-0.85)
Fixed Effects	Y, I	Y, I	Y, I	Y, I	Y, I	Y, I	Y, I	Y, I
R^2	0.445	0.465	0.446	0.464	0.451	0.433	0.451	0.429
N	8,417	1,962	8,417	1,962	8,763	2,258	8,763	2,402

Table 53
Minor Robustness Checks – Multivariate Results – continued

	Panel C Robustness Check Including Country-Fixed Effects			
	Model 1		Model 2	
	Unmatched Sample	Matched Sample	Unmatched Sample	Matched Sample
GENDER	-0.023 **			
	(-2.12)			
GENDERMATCH		0.006		
		(0.35)		
ABSGENDIFF			-0.001	-0.003
			(-1.06)	(-1.13)
NEGGAP			0.005	-0.009
			(-0.39)	(-0.32)
ABSGENDIFF × NEGGAP			0.003 *	0.000
			(1.75)	(0.21)
EPS	1.207 ***	1.271 ***	1.207 ***	1.273 ***
	(10.11)	(4.56)	(10.11)	(4.59)
ΔEPS	0.022	-0.007	0.022	-0.006
	(1.40)	(-0.34)	(1.39)	(-0.33)
Constant	0.137 ***	-0.050 ***	0.125 ***	-0.023
	(3.01)	(-0.64)	(2.74)	(-0.28)
Fixed Effects	Y, I, C	Y, I, C	Y,I,C	Y,I,C
R^2	0.457	0.439	0.457	0.439
N	8,872	2,258	8,872	2,258

Table 53 presents results of OLS Value Relevance regressions (Model 1) as well as the Gender Expectation Model (Model 2) for the matched and unmatched sample. Panel A presents results with a PSM caliper of 1 percent. Panel B presents results using lagged board data. Panel C presents results including country-fixed effects. Dependent variable is RET. RET is stock returns. GENDER is an indicator variable equal to 1 if a firm is included into the treatment group (has at least one female board member), 0 otherwise. GENDERMATCH is an indicator variable equal to 1 if a firm is included into the treatment group after applying the propensity score matching, 0 otherwise. ABSGENDIFF is the absolute value of GENDIFF. NEGGAP is an indicator variable equal to 1 if GENDIFF is negative, 0 otherwise. GENDIFF is the signed difference between the observed ratio of female board members in a given firm and the expected ratio of female board representation. EPS is earnings per share. ΔEPS is the year-to-year change in EPS. See Appendix A for detailed variable descriptions. Each model is estimated with year- (Y) and industry- (I) fixed effects. Standard errors are clustered at the firm-level (two-tailed t-statistics in parentheses). ***, **, and * denote significance at the 1 %, 5 %, and 10 % level, respectively.

6 Discussion

The primary research questions addressed in this study are (1) if female board participation has long-term effects on investors' perception of the firm and (2) what is the effect on a firm's stock performance if the actual degree of female board representation differs

from investors' expectations. Whereas most of the prior literature looks at the effects of board gender diversity on firms' accounting performance in single country settings (for a meta-analysis, cf. Post and Byron 2015), we focus on long-term capital market performance in an international setting. Our findings therefore offer some key contributions compared to prior literature.

First, our study offers two contributions with respect to theory. On the one hand, role congruity theory states that women may be perceived as less qualified for leadership positions (e.g., executive roles as board members) than men due to stereotypes and incongruent perceptions of social roles within a group. Thus, the theory has important implications for how stakeholder (e.g., investors) perceive female board members and how they may contribute to a firm's future profitability. Women on the board of directors are still relatively rare. Attributable to this, investors may believe that women do not have the same qualification and business experience as men. As a result, female board representation might be met with considerable scepticism. However, incongruent perceptions of board members based on gender differences would be captured in short-window stock reactions. Board members need some time to enact organizational changes and gain the trust of share- and stakeholder. Hence, the market requires an adjustment period to evaluate the actual quality of board directors instead of stereotypes. Therefore, looking at the longer-term effects of board gender diversity on stock performance is more convincing. Our results suggest that investors perceive (high-quality) female and male board members as being equivalent in the long-term, as we do not find any indication that stockholders' investment decisions seem to be influenced by directors' gender. However, while female board appointments may have short-window effects on firm value (e.g., Kang et al. 2010; Lee and James 2007; Schmid and Urban 2016), the market seemingly corrects this mispricing in the long run. Hence, this study contributes to a better understanding on how investors evaluate female board members despite of (alleged) perceived societal gender stereotypes. On the other hand, a lack of women on corporate boards might basically seem demonstrably unfair and could be perceived as a result of the 'glass ceiling'-effect (e.g., Arfken et al. 2004). Unsurprisingly, there is much pressure from various stakeholder groups to remedy this issue. Based on that, catering theory argues that firms cater to investors' and other stakeholders' demands by appointing more women to their boards (Ghosh et al. 2016). Given that an increasing number of investors demand a strive for more female board representation, firms might get 'punished' for non-compliance with

these expectations. However, we find no evidence that these firms systematically incur lower stock returns. Our results provide new evidence that board gender diversity does not come at the expense of investors – something opponents of mandatory gender quotas seem to commonly suggest (e.g., Ahern and Dittmar 2012). More importantly, our results also do not necessarily imply that the catering theory of board gender diversity is wrong. While some investors might consider gender diverse boards in their investment strategies (e.g., Byoun et al. 2016; Coffey and Fryxell 1991) there are others that apparently do not, indicating a balancing effect. Nevertheless, as there is evidence that an increasing number of investors explicitly considers gender equality in their investment decisions (e.g., Byoun et al. 2016; Rhode and Packel 2014), this result might shift over time.

Second, we offer some empirical contributions, as well. Hermalin and Weisbach (2003) state that corporate board composition and its association with performance is a profoundly endogenously issue. However, this factor is largely ignored by the far majority of previous studies. To mitigate threats associated with endogeneity, we employ the female director selection-model (Gul et al. 2011; Hillman et al. 2007) as a first step, in order to identify a comparable set of firms which only differs in terms of board gender diversity. As such, our PSM approach allows us to model a randomized quasi-natural experiment as closely as possible (Dehejia and Wahba 2002; Rosenbaum and Rubin 1983). In addition to the methodological contribution, most prior studies focus on single countries. Nonetheless, there is evidence that female board representation still differs significantly between countries (e.g., Grosvold 2011; Grosvold et al. 2016; Loy and Rupertus 2018a; Terjesen and Singh 2008), despite global advances to close the gender gap over recent years. Post and Byron (2015) argue that the relationship of female board representation and financial performance varies according to the level of gender parity in each country. Thus, one cannot generalize from prior single-country results. Hence, we test our hypotheses in a multi-country setting.

Third, our study contributes to the social and economic debate regarding female board representation. To this day, it is indisputable that female board representation should be increased due to social reasons such as fairness, equality, and participation. However, gender equality in the boardroom is no longer perceived solely as a social issue, but is also recommended based on economic arguments (i.e., the business case for women on board). Whereas one strand of literature finds that women have a positive impact on collaboration and effectiveness of corporate boards, another strand provides inconclusive

results with respect to the association of female board representation with corporate (accounting) performance. Having said that, to this day, the association of board gender diversity and long-term capital market performance is still unexplored. It is common sense that investors value 'high-quality' directors (e.g., Cao and Donnelly 2010; Schnatterly and Johnson 2014). Moreover, they emphasize board characteristics in their investment decisions (e.g., Bushee et al. 2014; Chung and Zhang 2011; Yermack 2006). However, according to our results, regardless of directors' gender. Given competing findings and (methodological) limitations of prior studies, we follow Rhode and Packel (2014) who argue that the business case of improved corporate financial performance through board gender diversity should neither be overstated nor generalized. Thus, we argue for a more nuanced position. When diversity is well-managed, it can improve decision-making and enhance a corporation's public image as it conveys a commitment to equal opportunity and inclusion (Radjavi 2012; Rhode and Packel 2014). Firms which adhere to the highest standards in terms of corporate governance – one of which undoubtedly is a commitment to gender equality – will enjoy positive capital market effects (e.g., Yermack 2006). Instead of political 'quick fixes' through mandatory gender quotas, stakeholders should encourage firms to continue to improve non-gender-biased hiring and promotion decisions to increase the pool of talented women in middle management and professional functions, who will eventually rise to top management and corporate boards. Hence, the focus in the current debate regarding the business case of female board representation should primarily be placed on arguments with respect to improvements of, among others, board-decision making, governance and strategy implementation, corporate reputation, and a firms' workforce rather than 'expecting' enhanced (accounting and market) performance. Gender diversity at the leadership level offers a strategic advantage in meeting the challenge of globalization, as boards will benefit from female leadership qualities, such as cross-cultural awareness and transformational leadership skills (Holton 2000; Terjesen and Singh 2008).

7 Limitations and Future Research

This research should be examined in light of some limitations. Our sample is based on Thomson Reuters' Asset 4 database which includes the largest public corporations across the globe. Therefore, our results might not be applicable to smaller stock corporations or smaller capital markets. Nonetheless, given the high market capitalization of our sample

constituents, we are confident that we appropriately cover board gender diversity's impact on an average investor's wealth. Moreover, some smaller economies with the highest gender equality (e.g., Norway or Sweden) are not included in our sample because there are nearly no firms without female board representation. However, these countries are predominantly subject to mandatory gender quotas throughout our sample period, and previous research suggests negative capital market effects of such quotas (Ahern and Dittmar 2012). Forcing an inclusion of these countries into our study would render our quasi-experimental approach useless.

Having said that, the distinct advantage of our PSM approach is that it allows us to draw quasi-causal inferences. However, while it seems to perform well in our international setting, there might be additional observable and unobservable factors which likewise affect director selection and female board representation. Future research may identify additional covariates and improve the precision of the estimation model.

Lastly, several countries in our sample have imposed mandatory gender quotas towards the end of – or after – our sample period (e.g., France, Germany). While we address this in a robustness check, future research may focus on the long-term capital market effects of quotas in these large economies.

8 Conclusion

In summary, our study sought to fill an important gap in the literature on investors' perception and long-term effects of board gender diversity on firms' capital market performance. Our results indicate that female board representation neither improves nor reduces firms' long-term stock performance if one appropriately controls for the endogeneity of corporate board appointments. Investors seem to perceive female and male board members as being equal in the long-term and do not base their investment decisions on directors' gender.

The board is the focal point of an organization's strategic decision-making. It monitors important day-to-day business activities, supervises management, liaises with auditors, (dis-)approves merger and financing decisions and, finally, selects and appoints top executives (e.g., Hambrick 2007; Hambrick and Mason 1984). Over recent years, there have been reinforced calls for increased board gender diversity to make their approaches to new solutions and strategies more heterogeneous (e.g., Adams and Kirchmaier 2015).

Unsurprisingly, many stakeholder and investor groups demand improvements on the individual firm level. Prior research attributes country-level differences in female board representation, among others, to differences in the percentage of women working full-time (Adams and Kirchmaier 2015), female enrolment in tertiary education, country-specific societal climates of gender equality (Loy and Rupertus 2018a), and differences in national political and institutional systems (e.g., Grosvold 2011; Grosvold and Brammer 2011; Grosvold et al. 2016; Terjesen and Singh 2008). Given the importance of this topic, it seems necessary to go beyond the conventional thinking in terms of the business case for gender diversity and broaden the perspective to also incorporate social and ethical aspects. As such, our study might be the missing bridge between both perspectives.

Appendix A: Variable Descriptions

Variable	Description
ABSGENDIFF	The absolute difference between the observed proportion of female board members in a given firm and the expected ratio of female board representation. The expected ratio in the main analysis is the average ratio of female board representation across all firms in a given industry, by country and year.
AGE	The natural logarithm of firm age defined as the number of years since the date of incorporation (Source: Worldscope).
BHR	The average buy-and-hold return over the preceding three years (Source: Datastream).
BOARD_POL	Indicator variable equal to 1 if the company has a policy regarding the gender diversity of its board, and 0 otherwise (Source: Asset4; code: CGBSDP0013).
BOUTSIDE	The natural logarithm of outside directors measured as the difference between total number of board members (Source: Asset4; code: CGBSDP060) and the number of executive directors.
CORPCODE	Indicator variable equal to 1 if the country has a voluntary corporate governance code stipulation regarding gender in year t, 0 otherwise (Source: Loy and Rupertus 2018a; Terjesen et al. 2015).
EPS	Earnings per share (i.e., net income before extraordinary items scaled by the number of shares outstanding; Source: Worldscope).
ΔEPS	The year-to-year change in earnings per share (EPS) scaled by the stock price at the beginning of the fiscal year (Source: Worldscope).
GENDER	Indicator variable equal to 1 if a firm is included into the treatment group, 0 otherwise. Our treatment group consists of firm-years, which exhibit at least one female board member (Source: Asset4; code: CGBSO17V).
GENDERMATCH	Indicator variable equal to 1 if a firm is (still) included into the treatment group after applying propensity score matching, and 0 otherwise.
GENDIFF	The signed difference between the observed proportion of female board members in a given firm and the expected ratio of female board representation.
GENQUOT	Observed percentage of women on the boards of directors (Source: Asset4; code: CGBSO17V).
GGGI	Index value of the Global Gender Gap Index (Source: World Economic Forum).
IAR	Industry-adjusted stock returns estimated by the difference between the average stock return in a given industry and a firms' raw stock return (Source: Datastream).
IND	Indicator variables for industry-fixed effects based on the Fama/French 12 industry portfolio (Source: Worldscope).
LEV	Ratio of total debt to total shareholders' equity (Source: Worldscope).

Variable	Description
MAR	Market-adjusted stock returns estimated by the difference between the average market return of the top index of each country and a firm's raw stock return.
NEGGAP	Indicator variable equal to 1 if GENDIFF is negative, 0 otherwise.
QUOTA	Indicator variable equal to 1 if the country has a mandatory gender quota in year t, and 0 otherwise (Source: Loy and Rupertus 2018a; Terjesen et al. 2015).
RET	Stock return is computed as year-end share price plus dividends per share minus prior year-end share price divided by prior year-end share price. We employ the period from nine months before to three months after fiscal year-end. This corresponds with the timing of financial statement disclosures and shareholders' meetings (Source: Datastream).
RISK	Rolling five year-standard deviation of cash flow from operations (Source: Worldscope).
ROA	Return on assets is computed as net income before extraordinary items divided by total assets at fiscal year-end (Source: Worldscope).
SIZE	Firm size is proxied by the natural logarithm of total assets (Source: Worldscope).
SGROWTH	Average sales growth over the prior three fiscal years (Source: Worldscope).
TD	Total diversification computed as $_{i=1} = s_i \times \ln(\frac{1}{s_i})$, where s_i is the share of the ith industry segment compared to total firm sales (Source: Palepu 1985).
YEAR	Indicator variables for year-fixed effects.

Appendix B: Country-by-Country Matching Model Results

	Australia		Japan		U.K.		U.S.	
	Unmatched Sample	Matched Sample	Unmatched Sample	Matched Sample	Unmatched Sample	Matched Sample	Unmatched Sample	Matched Sample
	GENDER	GENDER-MATCH	GENDER	GENDER-MATCH	GENDER	GENDER-MATCH	GENDER	GENDER-MATCH
ROA	-2.130 ***	-0.451	5.193 **	-1.756	1.805	-0.891	0.206	-0.883
	(-1.56)	(-0.27)	(2.33)	(-0.59)	(0.78)	(-0.28)	(0.26)	(-0.94)
SIZE	0.976 ***	-0.216	0.486 ***	0.050	0.431 ***	-0.253	0.357 ***	-0.292 ***
	(6.35)	(-1.03)	(6.16)	(0.47)	(3.06)	(-1.29)	(5.95)	(-3.50)
BOUTSIDE	0.197	-0.117	0.015	-0.563 **	0.886 ***	-0.333	2.191 ***	-0.477 **
	(0.77)	(-0.39)	(0.10)	(-2.43)	(2.94)	(-0.86)	(14.11)	(-2.42)
RISK	-3.053	-0.351	0.534	-0.765	-11.83 ***	1.729	-0.488	0.152
	(-1.23)	(-0.13)	(0.14)	(-0.16)	(-2.91)	(0.33)	(-0.28)	(0.08)
TD	0.857 **	0.017	0.101	-0.008	0.586	-0.241	0.205	-0.195
	(2.46)	(0.04)	(0.67)	(-0.04)	(1.51)	(-0.48)	(1.49)	(-1.16)
AGE	0.531 ***	-0.286	0.303 **	0.072	-0.246	0.183	0.292 ***	-0.168 *
	(2.86)	(-1.20)	(2.42)	(0.43)	(-1.42)	(0.81)	(3.97)	(-1.78)
SGROWTH	-0.305	0.096	2.636 ***	0.809	-1.848 **	-0.684	-1.846 ***	0.023
	(-1.14)	(0.29)	(3.41)	(0.77)	(-2.15)	(-0.61)	(-4.48)	(0.05)
LEV	-0.329 ***	0.184	-0.026	0.013	0.024	0.021	-0.013	0.009
	(-3.49)	(1.09)	(-0.80)	(0.27)	(1.59)	(1.05)	(-0.95)	(0.54)
Constant	-14.09 ***	3.756	-13.56 ***	0.529	-4.379 **	2.013	-8.512 ***	5.699 ***
	(-6.65)	(1.26)	(-7.88)	(0.24)	(-2.10)	(0.71)	(-9.18)	(4.22)
Fixed Effects	Y, I	Y, I	Y, I	Y, I	Y, I	Y, I	Y, I	Y, I
R^2	0.325	0.022	0.103	0.028	0.254	0.046	0.245	0.031
N	512	202	2,272	488	358	157	3,692	896

Appendix B: Country-by-Country Matching Model Results – continued

	Brazil		Canada		France		Germany	
	Unmatched Sample	Matched Sample	Unmatched Sample	Matched Sample	Unmatched Sample	Matched Sample	Unmatched Sample	Matched Sample
	GENDER	GENDER-MATCH	GENDER	GENDER-MATCH	GENDER	GENDER-MATCH	GENDER	GENDER-MATCH
ROA	13.85	-22.19	0.078	3.637	-5.014	14.37	5.591	1.937
	(1.58)	(-0.89)	(0.02)	(0.85)	(-0.84)	(1.28)	(1.29)	(0.33)
SIZE	-0.790 *	1.765	0.875 ***	-0.068	0.613 *	0.903	-0.251	0.266
	(-1.94)	(1.01)	(4.32)	(-0.25)	(1.82)	(1.38)	(-1.23)	(0.91)
BOUTSIDE	-0.383	-1.492	2.140 ***	-0.283	0.916	-2.473 **	2.815 ***	-0.469
	(-0.50)	(-1.07)	(4.07)	(-0.38)	(1.43)	(-2.03)	(5.13)	(-0.63)
RISK	1.599	5.084	-18.74 ***	-1.201	-23.62 **	20.30	-4.699	10.69
	(0.24)	(0.47)	(-2.72)	(-0.12)	(-2.09)	(1.08)	(-0.62)	(0.91)
TD	1.316 *	-6.058 *	1.668 ***	-0.106	0.092	-5.216 ***	-0.510	-0.605
	(1.66)	(-1.68)	(3.44)	(-0.14)	(0.12)	(-2.71)	(-1.04)	(-0.84)
AGE	-0.243	1.443 *	-0.303	-0.389	0.783 *	0.249	-0.912 ***	-0.072
	(-0.56)	(1.69)	(-0.92)	(-0.76)	(1.71)	(0.30)	(-3.38)	(-0.20)
SGROWTH	-2.682	-2.034	-0.117	0.442	-0.032	0.398	4.791 *	-4.145
	(-1.14)	(-0.25)	(-0.10)	(0.25)	(-0.02)	(0.17)	(1.81)	(-1.23)
LEV	0.472 ***	-1.008	0.006	-0.047	-0.118 **	-0.009	-0.028	0.021
	(2.62)	(-1.58)	(0.09)	(-0.49)	(-2.14)	(-0.10)	(-0.79)	(0.51)
Constant	9.999	-21.96	-15.82 ***	3.323	-10.38 **	-8.070	4.940	-3.536
	(1.55)	(-0.78)	(-5.75)	(0.86)	(-2.03)	(-0.81)	(1.48)	(-0.75)
Fixed Effects	Y, I	Y, I	Y, I	Y, I	Y, I	Y, I	Y, I	Y, I
R^2	0.463	0.365	0.539	0.058	0.319	0.214	0.349	0.048
N	94	34	422	84	175	54	292	100

Appendix B: Country-by-Country Matching Model Results – continued

	India		Italy		Singapore		Spain	
	Unmatched Sample	Matched Sample	Unmatched Sample	Matched Sample	Unmatched Sample	Matched Sample	Unmatched Sample	Matched Sample
	GENDER	GENDER-MATCH	GENDER	GENDER-MATCH	GENDER	GENDER-MATCH	GENDER	GENDER-MATCH
ROA	-8.998	0.466	-15.19	49.44	3.932	-6.711	6.385	33.28
	(-1.47)	(0.05)	(-0.95)	(1.34)	(0.51)	(-0.66)	(1.12)	(0.56)
SIZE	0.209	-0.283	0.313	-0.953	1.140 **	-0.762	1.236 *	-5.143
	(0.65)	(-0.62)	(1.08)	(-1.58)	(2.29)	(-1.09)	(1.65)	(-1.14)
BOUTSIDE	2.810 ***	-0.131	-0.787	5.054 *	1.578	1.300	2.413 **	-6.784
	(4.06)	(-0.14)	(-0.61)	(1.67)	(1.47)	(1.08)	(2.03)	(-1.43)
RISK	13.76	-15.82	-5.660	12.13	2.758	-3.756	-3.402	289.28 **
	(1.17)	(-0.62)	(-1.05)	(1.29)	(0.35)	(-0.37)	(-0.43)	(1.99)
TD	-0.035	0.222	-1.628	-3.638 *	-2.588 ***	1.205	-1.235	-7.654
	(-0.07)	(0.31)	(-1.57)	(-1.74)	(-3.24)	(1.04)	(-0.97)	(-1.38)
AGE	-0.824	0.492	0.599	1.959 *	-0.985 **	0.443	-0.440	7.349
	(-1.53)	(0.55)	(1.31)	(1.87)	(-2.02)	(0.71)	(-0.33)	(1.16)
SGROWTH	-0.926	5.970	-0.043	8.864 **	2.620	-2.478	-2.983	5.971
	(-0.75)	(1.35)	(-0.03)	(1.98)	(1.23)	(-0.98)	(-0.80)	(0.50)
LEV	-0.099	-0.102	-0.037	-0.093	-0.198	0.218	-0.071	0.222
	(-1.05)	(-0.58)	(-0.54)	(-0.52)	(-1.27)	(0.99)	(-0.69)	(0.43)
Constant	-6.669	4.023	-4.483	-2.905	-11.319	5.726	-19.29	67.035
	(-1.09)	(0.44)	(-0.92)	(-0.29)	(-1.54)	(0.60)	(-1.22)	(0.79)
Fixed Effects	Y, I	Y, I	Y, I	Y, I	Y, I	Y, I	Y, I	Y, I
R^2	0.339	0.088	0.223	0.242	0.242	0.084	0.419	0.480
N	154	67	79	35	111	58	103	25

Appendix B: Country-by-Country Matching Model Results – continued

	Switzerland	
	Unmatched Sample	Matched Sample
	GENDER	GENDER-MATCH
ROA	-8.998	0.466
	(-1.47)	(0.05)
SIZE	0.209	-0.283
	(0.65)	(-0.62)
BOUTSIDE	2.810 ***	-0.131
	(4.06)	(-0.14)
RISK	13.76	-15.82
	(1.17)	(-0.62)
TD	-0.035	0.222
	(-0.07)	(0.31)
AGE	-0.824	0.492
	(-1.53)	(0.55)
SGROWTH	-0.926	5.970
	(-0.75)	(1.35)
LEV	-0.099	-0.102
	(-1.05)	(-0.58)
Constant	-6.669	4.023
	(-1.09)	(0.44)
Fixed Effects	Y, I	Y, I
R^2	0.339	0.088
N	154	67

Appendix B presents results of logit models for the prediction of female board representation (Equation 8) for each country in our sample. Dependent variable is GENDER or GENDER-MATCH. GENDER is an indicator variable equal to 1 if a firm is included into the treatment group (has at least one female board member), 0 otherwise. GENDERMATCH is an indicator variable equal to 1 if a firm is included into the treatment group after applying the PSM model, 0 otherwise. ROA is return on assets. SIZE is the natural logarithm of total assets. BOUTSIDE is the natural logarithm of outside directors. RISK is the standard deviation of cash flows from operations. TD is total diversification. AGE is the natural logarithm of firm age. SGROWTH is the average sales growth over the prior three fiscal years. LEV is the ratio of total debt to share-holders' equity. See Appendix A for detailed variable descriptions. Standard errors are clustered on the firm-level (two-tailed t-statistics in parentheses). Each model is estimated with year- (Y) and industry- (I) fixed effects. ***, **, and * denote significance at the 1 %, 5 %, and 10 % level, respectively.

CONCLUDING REMARKS

This dissertation consists of two chapters on disclosure regulation (i.e., CAP disclosures) and two chapters on corporate governance mechanisms (i.e., board diversity). Chapter 1 is the first comprehensive study on the regulatory framework, prior research, and the implementation in practice of CAP disclosures over a 16-year period. It further shows how related disclosures comply with the SEC guidelines. Chapter 2 focus on whether and how CAP disclosures reflect instances of measurement uncertainties embedded in individual financial statement positions and how accruals flagged as CAPs are related to future cash flows. Chapter 3 expands existing knowledge in terms of supply and demand-side factors associated with considerable cross-country variation in female board representation. Chapter 4 explores the association of board gender diversity with long-term effects on shareholders' wealth. Each chapter offers some key contributions with respect to theory, methodology, and the current public debate regarding disclosure regulation and governance mechanisms, which I will discuss thereafter. I focus on the key contributions of my dissertation. Further contributions and future research suggestions are described in detail in each chapter.

First, my findings offer theoretical contribution. With respect to Chapter 2 of this dissertation, it remains unclear whether CAPs reflect instances of measurement uncertainties in accrual-based measures and how they are related with future cash flows. Analyzing the predictive value of accruals that are flagged as CAPs is far from straightforward. Firms may see CAPs as a mere 'compliance exercise' without providing any real information content. Further, implementing CAPs in the regulators intended manner could have two contrary effects. First, if management correctly identifies measurement uncertainties in their accounting estimates, they might consider such information in the measurement process of financial statements, thus, leading to improved or not to diminished predictive ability with respect to future cash flows. Second, uncertainties in the measurement process result in incomplete accounting information and imprecise valuation inputs. Consequently, management cannot make reliable and accurate valuations, which might lead to reduced predictive power of accruals with respect to future cash flows. According to our results, we document that accruals flagged as CAPs are less persistent with respect to future cash flows. Thus, our results are in line with the studies of Richardson et al. (2005) and Sloan (1996) and show that subjectivity and uncertainties embedded in the measurement process of accrual-based measures diminish their usefulness to predict future cash

flows. As there is no prior evidence on how to identify subjective and uncertain accruals, we demonstrate that CAP disclosures provide such information.

The second part of this dissertation provides theoretical contribution about investors' perception of women on corporate boards. In this context, role congruity theory provides important implications on how stakeholders perceive female board members. Because the number of female board members is still relatively low, one might argue that investors believe women do not have the same qualification as men, and thus might be viewed with considerable skepticism. While prior studies solely focus on short-window effects of female board members on firm value, we extend prior evidence on role congruity theory and its implications for how investors perceive females on corporate boards in the long-term. Furthermore, catering theory argue that firms cater to investors' demands by appointing more women to the boards. Given that an increasing number of investors demand female board representation, firms might get 'punished' for non-compliance with these expectations. We find no evidence that firms incur lower stock returns if they fail to do so. Nonetheless, our results do not imply that catering theory in the context of board gender diversity is wrong. We argue that while there is empirical evidence that some investors explicitly consider gender equality in their investment decisions (e.g., Rhode and Packel 2014), others do not, indicating a balancing effect.

Second, this dissertation offers methodological and empirical contributions. Regarding the first part of this dissertation, to best of my knowledge, I am the first who focus on (qualitative and descriptive) reporting characteristics of CAP disclosures. While quantitative information about CAPs have been studied before, prior research neglected the qualitative information content beyond CAP disclosures. This is quite interesting because regulators and practitioners repeatedly emphasized the important to provide clear, understandable, and simple disclosures. In recent years, the importance of textual analysis in accounting research has grown tremendously (e.g., Hering 2018; Loughran and McDonald 2016), because it allows to abstract and characterize narrative information across several distinct dimensions. Using tools from computational linguistics, I provide first evidence about four reporting characteristics that are in line with the evolving textual analysis literature. Specifically, I find that the average length of CAPs has constantly increased from 2001 to 2016, CAP disclosures are very similar over time, they do not contain a large portion of specific information, and are highly complex (i.e., difficult to read). It

seems that my findings stand in contrast with the SEC intention to provide understandable, clear, and specific disclosures in plain English. However, as my findings here are only descriptive in nature, they need to be interpreted with caution. Nonetheless, my findings provide initial evidence on the reporting characteristics beyond CAPs. Future research could analyze how CAP disclosures are associated with a firm's information environment or whether and how high quality CAP disclosures help analysts to improve their forecasts about future cash flows and earnings.

The second part of this dissertation focuses on the recent governance research call to address unobserved heterogeneity and simultaneity in board governance choices. This indicates that current values of governance variables are dynamically related to its historical values (Wintoki et al. 2012). As such, we employ dynamic panel estimation (SYS-GMM) in Chapter 3 (Blundell and Bond 1998) and use lagged values of the explanatory variable (i.e., lagged values of the average ratio of female board members) as instruments for current changes of the same variable to alleviate concerns of autocorrelation (e.g., Wintoki et al. 2012). Furthermore, corporate board diversity and its association with performance is a profoundly endogenously issue (Hermalin and Weisbach 2003). Accordingly, it cannot be said with certainty whether performance (e.g., capital market performance) drives board diversity or whether board diversity is the outcome of other unobservable variables, which are also related to performance. Interestingly, the issue of endogeneity is largely ignored by the majority of previous studies. Therefore, we implement a quasi-experimental approach to mitigate threats of endogeneity in Chapter 4. Specifically, we employ propensity score matching (PSM) to identify a comparable set of firms, which only differ in terms of board gender diversity.

Third, my dissertation contributes to the current public debate regarding CAP disclosures as well as board gender diversity. The SEC started a series of initiatives to make corporate disclosures more transparent. However, stakeholders have raised concerns about the usefulness of corporate disclosures in recent years due an increasing number of new and additional financial reporting requirements as well as voluntary disclosures made by firms (the so-called '*information overload*') (e.g., Hellman et al. 2018; Schick et al. 1990). Thus, it is questionable whether new and/or additional disclosure regulations are helpful for financial statement users. It is important that firms provide convincing information about financial statements positions that are inherently more uncertain, so that financial statement users can incorporate such information into their decisions. We shed light on

the role of CAP to provide such information. The conclusions drawn from Chapter 1 and 2 show that CAPs fulfil the SEC's goal to communicate highly uncertain accounting policies and estimates. Further, they help investors to obtain greater insights into the variation of measurement uncertainties and the way uncertainties inherent in accounting estimates affect a company's financial performance. Thus, we contribute to the debate regarding corporate disclosures about measurement uncertainties in analyzing whether CAPs provide information about the credibility of financial statements. Furthermore, the IASB is currently discussing on how to improve disclosures about accounting policies, judgements, and estimation uncertainties, because there is increasing concern that current requirements do not fulfil their intended purpose (IASB 2019). While U.S. firms provide consistent and extensive information about their CAPs in the MD&A, disclosure practice about estimation uncertainties in the notes to the financial statement is mixed (Fülbier et al. 2017). Based on my findings as well as prior evidence, I suggest that the IASB might consider parts of the CAP regulation in their future discussion to revise and improve accounting policy disclosures.

Moreover, the motives to appoint women to the board of directors are based on social as well as business-related theories and arguments (i.e., the social and business case of women on boards, cf. Figure 9). Whereas it is indisputable, that female board representation should be increased because of social reasons, prior literature finds inconclusive results with respect to the association of female board representation and corporate (accounting) performance. It further neglects the relation to long-term market performance. Our results expand existing literature suggesting that investors perceive female and male board members as being equivalent in the long-term. It is common sense that investors value 'high quality' directors and emphasize board characteristics in their investment decisions. However, our results show that this is unrelated to a director's gender. We argue that the business case of board gender diversity should not be overstated nor generalized. Specifically, we argue that firms which adhere to the highest standards of corporate governance – one of which seems to be the commitment to gender equality – will enjoy positive effects on capital markets (e.g., Yermack 2006). Based on our findings in Chapter 3 and 4, we conclude that legislators might rather focus on supply-side measures and firms might improve their non-gender-biased hiring and promotion decisions to give women more opportunities to move into managerial and professional roles and thus, to increase the pool of talented women for top management positions and corporate boards.

Given the importance of this topic, it seems necessary to go beyond the conventional thinking of the business case regarding female board representation and, in particular, incorporate social and ethical aspects.

Overall, this doctoral dissertation contributes to the emerging field of literature examining disclosure regulation (i.e., CAP disclosures) and corporate governance mechanisms (i.e., board diversity). Given the limitations of the four studies presented, I provide several fruitful avenues for future studies. As regulators, practitioners, politicians, and the international press still intensively discuss both topics, I am excited about seeing future research in these fields to enhance our knowledge on disclosure regulation as well as governance mechanisms.

References

AbbVie Inc. (2016). Annual Report (10-K), AbbVie Inc. Available online: https://www.sec.gov/Archives/edgar/data/1551152/000155115217000004/abbv-12312016x10k.htm (accessed 19 March 2019).

Abrams, D., Thomas, J. and Hogg, M. A. (1990). Numerical distinctiveness, social identity and gender salience. *British Journal of Social Psychology*, 29(1), 87–92.

Adams, R. B. and Ferreira, D. (2009). Women in the Boardroom and their impact on Governance and Performance. *Journal of Financial Economics*, 94(2), 291–309.

Adams, R. B. and Funk, P. (2012). Beyond the Glass Ceiling: Does Gender Matter? *Management Science*, 58(2), 219–235.

Adams, R.B. and Kirchmaier, T. (2015). Barriers to Boardroom. *ECGI - Finance Working Paper No. 347/2013, Asian Finance Association (AsFA) 2013*, University of New South Wales and University of Manchester. Available online: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2192918 (accessed 19 March 2019).

Ahern, K. R. and Dittmar, A. K. (2012). The Changing of the Boards: The Impact on Firm Valuation of Mandated Female Board Representation. *The Quarterly Journal of Economics*, 127(1), 137–197.

AICPA (1994). Improving Business Reporting - A Customer Focus. The Jenkins Report, American Institute of Chartered Public Accountants, New York. Available online: https://www.aicpa.org/content/dam/aicpa/interestareas/frc/accountingfinancialreporting (accessed 19 March 2019).

Alchian, A. A. and Demsetz, H. (1972). Production, Information Costs, and Economic Organization. *The American Economic Review*, 62(52), 777–795.

Allen, E. J., Larson, C. R. and Sloan, R. G. (2013). Accrual reversals, earnings and stock returns. *Journal of Accounting and Economics*, 56(1), 113–129.

Alphabet Inc. (2016). Annual Report (10-K), Alphabet Inc. Available online: https://www.sec.gov/Archives/edgar/data/1652044/000165204417000008/goog10-kq42016.htm (accessed 19 March 2019).

APB (1972). Disclosure of Accounting Policies, Accounting Principles Board. Available online: https://www.fasb.org/jsp/FASB/Document_C/DocumentPage?cid=1176156418903&acceptedDisclaimer=true (accessed 19 March 2019).

Arfken, D. E., Bellar, S. L. and Helms, M. M. (2004). The Ultimate Glass Ceiling Revisited: The Presence of Women on Corporate Boards. *Journal of Business Ethics*, 50(2), 177–186.

Barckow, A. (2018). Regulation in Corporate Disclosure. *Schmalenbach Business Review*, 1–5.

Barth, M. E. (2006). Including estimates of the future in today's financial statements. *Accounting Horizons*, 20(3), 271–285.

Barth, M. E., Clinch, G. and Israeli, D. (2016). What do accruals tell us about future cash flows? *Review of Accounting Studies*, 21(3), 768–807.

Barth, M. E., Cram, D. P. and Nelson, K. K. (2001). Accruals and the prediction of future cash flows. *The Accounting Review*, 76(1), 27–58.

Barth, M. E. and Schipper, K. (2008). Financial Reporting Transparency. *Journal of Accounting, Auditing & Finance*, 23(2), 173–190.

Bauman, M. P. and Shaw, K. W. (2014). An Analysis of Critical Accounting Estimate Disclosures of Pension Assumptions. *Accounting Horizons*, 28(4), 819–845.

Bear, S., Rahman, N. and Post, C. (2010). The Impact of Board Diversity and Gender Composition on Corporate Social Responsibility and Firm Reputation. *Journal of Business Ethics*, 97(2), 207–221.

Beaver, W. H. (1991). Problems and paradoxes in the financial reporting of future. *Accounting Horizons*, 5(4), 122–134.

Berelson, B. (1952). Content Analysis in Communication Research, Free Press, New York.

Billings, M. B. (2011). Discussion of "Critical Accounting Policy Disclosures". *Journal of Accounting, Auditing & Finance*, 77–80.

Bird, R., Reddy, K. and Yeung, D. (2014). The relationship between uncertainty and the market reaction to information: Is it influenced by stock-specific characteristics? *International Journal of Behavioral Accounting and Finance*, 4(2), 113–132.

Blair, M. M. (1995). Rethinking Assumptions behind Corporate Governance. *Challenge*, 38(6), 12–17.

Bloomfield, R. J. (2002). The 'Incomplete Revelation Hypothesis' and Financial Reporting. *Accounting Horizons*, 16(3), 233–243.

Blundell, R. and Bond, S. (1998). Initial conditions and moment restrictions in dynamic panel data models. *Journal of Econometrics*, 87(1), 115–143.

Boulouta, I. (2013). Hidden Connections: The Link Between Board Gender Diversity and Corporate Social Performance. *Journal of Business Ethics*, 113(2), 185–197.

Bozanic, Z. and Thevenot, M. (2015). Qualitative Disclosure and Changes in Sell-Side Financial Analysts' Information Environment. *Contemporary Accounting Research*, 32(4), 1595–1616.

Brammer, S., Millington, A. and Pavelin, S. (2007). Gender and Ethnic Diversity Among UK Corporate Boards. *Corporate Governance: An International Review*, 15(2), 393–403.

Brown, S. V. and Tucker, J. W. (2011). Large-Sample Evidence on Firms' Year-over-Year MD&A Modifications. *Journal of Accounting Research*, 49(2), 309–346.

Burke, R. J. (1997), *Women on Corporate Boards of Directors: A Needed Resource*, in Burke, R.J. (Ed.), Women in Corporate Management, Springer, Dordrecht, 37–43.

Bushee, B. J., Carter, M. E. and Gerakos, J. (2014). Institutional Investor Preferences for Corporate Governance Mechanisms. *Journal of Management Accounting Research*, 26(2), 123–149.

Bushman, R. M., Piotroski, J. D. and Smith, A. J. (2004). What Determines Corporate Transparency? *Journal of Accounting Research*, 42(2), 207–252.

Bushman, R. M. and Smith, A. J. (2003). Transparency, Financial Accounting Information, and Corporate Governance. *Economic Policy Review*, 9(1), 65–87.

Byoun, S., Chang, K. and Kim, Y. S. (2016). Does Corporate Board Diversity Affect Corporate Payout Policy? *Asia-Pacific Journal of Financial Studies*, 45(1), 48–101.

Cadbury, A. (1992). Report of the Committee on the Financial Aspects of Corporate Governance, Gee, London.

California Public Employees' Retirement System (2016). Comment Letter - Concept Release: Business and Financial Disclosure Required by Regulation S-K, California Public Employees' Retirement System. Available online: https://www.sec.gov/comments/s7-06-16/s70616-267.pdf (accessed 19 March 2019).

Campbell, K. and Minguez Vera, A. (2010). Female board appointments and firm valuation: short and long-term effects. *Journal of Management & Governance*, 14(1), 37–59.

Campbell, K., Sefcik, S. E. and Soderstrom, N. S. (2003). Disclosure of private information and reduction of uncertainty: Environmental liabilities in the chemical industry. *Review of Quantitative Finance and Accounting*, 21(4), 349–378.

Cao, T. and Donnelly, R. (2010). Active institutional investors and the board of directors. *International Journal of Business Governance and Ethics*, 5(4), 301–322.

Carter, D. A., D'Souza, F., Simkins, B. J. and Simpson, W. G. (2010). The Gender and Ethnic Diversity of US Boards and Board Committees and Firm Financial Performance. *Corporate Governance: An International Review*, 18(5), 396–414.

Carter, D. A., Sumkins, B. J. and Simpson, W. G. (2003). Corporate Governance, Board Diversity and Firm Value. *The Financial Review*, 38(1), 33–53.

Cassell, C. A., Dreher, L. M. and Myers, L. A. (2013). Reviewing the SEC's Review Process: 10-K Comment Letters and the Cost of Remediation. *The Accounting Review*, 88(6), 1875–1908.

Cejka, M. A. and Eagly, A. H. (1999). Gender-Stereotypic Images of Occupations Correspond to the Sex Segregation of Employment. *Personality and Social Psychology Bulletin*, 25(4), 413–423.

Center for Audit Quality (2016). Comment Letter - Concept Release: Business and Financial Disclosure Required by Regulation S-K, Center for Audit Quality. Available online: https://www.sec.gov/comments/s7-06-16/s70616-216.pdf (accessed 19 March 2019).

Chartered Professional Accountants Canada (2016). Corporate Reporting Briefing: What should be disclosed about estimation uncertainty?, Chartered Professional Accountants, Toronto, CA. Available online: https://www.cpacanada.ca/en/business-and-accounting-resources/financial-and-non-financial-reporting/mdanda-and-other-financial-reporting/publications/estimation-uncertainty-what-should-be-disclosed (accessed 19 March 2019).

Chen, H.(J.), Chen, J.V. and Li, F. (2019). Do the Notes to the Financial Statements Provide Information about the Degree of Subjectivity in Accruals? *Working Paper*, Tsinghua University - PBC School of Finance, University of Illinois at Chicago and Shanghai Jiaotong University. Available online: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2738842 (accessed 1 March 2019).

Chen, S., Ni, X. and Tong, J. Y. (2016). Gender Diversity in the Boardroom and Risk Management: A Case of R&D Investment. *Journal of Business Ethics*, 136(3), 599–621.

Chevron Corporation (2016). Comment Letter - Concept Release: Business and Financial Disclosure Required by Regulation S-K, Chevron Corporation. Available

online: https://www.sec.gov/comments/s7-06-16/s70616-315.pdf (accessed 19 March 2019).

Cho, J.S., Park, J. and Warfield, T.D. (2005). Critical Accounting Policy and Estimate Disclosures. *Working Paper*, University of Wisconsin and University of Seoul.

Christensen, B. E., Glover, S. M. and Wood, D. A. (2012). Extreme estimation uncertainty in fair value estimates: Implications for Audit Assurance. *AUDITING: A Journal of Practice & Theory*, 31(1), 127–146.

Christensen, B.E., Glover, S.M., Omer, T.C. and Shelley, M.K. (2014). Does estimation uncertainty affect investors' preference for the form of financial statement presentation? *Working Paper*, Texas A&M University, Brigham Young University and University of Nebraska-Lincoln. Available online: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2163878 (accessed 19 March 2019).

Chung, K. H. and Zhang, H. (2011). Corporate Governance and Institutional Ownership. *The Journal of Financial and Quantitative Analysis*, 46(1), 247–273.

Chychyla, R., Leone, A. J. and Minutti-Meza, M. (2018). Complexity of financial reporting standards and accounting expertise. *Journal of Accounting and Economics*, 67(1), 226–253.

Coase, R. H. (1937). The Nature of the Firm. *Economica*, 4(16), 386–405.

Coffey, B. S. and Fryxell, G. E. (1991). Institutional ownership of stock and dimensions of corporate social performance: An empirical examination. *Journal of Business Ethics*, 10(6), 437–444.

Cole, C. J. and Jones, C. L. (2005). Management discussion and analysis: A review and implications of future research. *Journal of Accounting Literature*, 24, 135–174.

Cumming, D., Leung, T. Y. and Rui, O. (2015). Gender Diversity and Securities Fraud. *Academy of Management Journal*, 58(5), 1572–1593.

Dechow, P. M. (1994). Accounting earnings and cash flows as measures of firm performance. *Journal of Accounting and Economics*, 18(1), 3–42.

Dechow, P. M. and Dichev, I. D. (2002). The quality of accruals and earnings: The role of accrual estimation errors. *The Accounting Review*, 77(s-1), 35–59.

Dechow, P. M. and Ge, W. (2006). The persistence of earnings and cash flows and the role of special items: Implications for the accrual anomaly. *Review of Accounting Studies*, 11(2-3), 253–296.

Dechow, P. M., Kothari, S. P. and L. Watts, R. (1998). The relation between earnings and cash flows. *Journal of Accounting and Economics*, 25(2), 133–168.

Dehejia, R. H. and Wahba, S. (2002). Propensity Score-Matching Methods for Nonexperimental Causal Studies. *Review of Economics and Statistics*, 84(1), 151–161.

Deloitte (2017). Women in the boardroom - A global perspective, Deloitte. Available online: https://www2.deloitte.com/content/dam/Deloitte/de/Documents/audit/Women%20in%20the%20boardroom%20a%20global%20perspective%20fifth%20edition.pdf (accessed 19 March 2019).

Demb, A. and Neubauer, F.-F. (1992). The corporate board: Confronting the paradoxes. *Long Range Planning*, 25(3), 9–20.

Dezsö, C. L. and Ross, D. G. (2012). Does female representation in top management improve firm performance? A panel data investigation. *Strategic Management Journal*, 33(9), 1072–1089.

Dobbin, F. and Jung, J. (2011). Corporate Board Gender Diversity and Stock Performance: The Competence Gap or Institutional Investor Bias. *North Carolina Law Review*, 89(3), 809–838.

Donaldson, L. and Davis, J. H. (1991). Stewardship Theory or Agency Theory: CEO Governance and Shareholder Returns. *Australian Journal of Management*, 16(1), 49–64.

Duffie, D. and Lando, D. (2001). Term structures of credit spreads with incomplete accounting information. *Econometrica*, 69(3), 633–664.

Eagly, A. H. and Karau, S. J. (2002). Role congruity theory of prejudice toward female leaders. *Psychological Review*, 109(3), 573–598.

Easley, D. and O'Hara, M. (2004). Information and the Cost of Capital. *The Journal of Finance*, 59(4), 1553–1583.

Eilifsen, A., Hamilton, E.L. and Messier, W.F. (2017). The importance of quantifying uncertainty: Examining the effects of sensitivity analysis and audit materiality disclosures on investors' judgments and decisions. *Working Paper*, NHH Norwegian School of Economics and University of Nevada. Available online: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2966291 (accessed 19 March 2019).

Elo, S. and Kyngäs, H. (2008). The qualitative content analysis process. *Journal of Advanced Nursing*, 62(1), 107–115.

El-Sayed Ebaid, I. (2011). Accruals and the prediction of future cash flows. *Management Research Review*, 34(7), 838–853.

Erhardt, N. L., Werbel, J. D. and Shrader, C. B. (2003). Board of Director Diversity and Firm Financial Performance. *Corporate Governance*, 11(2), 102–111.

European Commission (2016). Gender Balance on Corporate Boards - Europe is cracking the glass ceiling, Brussels. Available online: http://ec.europa.eu/newsroom/document.cfm?doc_id=46280 (accessed 19 March 2019).

ExplosionAI (2019). SpaCy - Industrial-Strength Natural Language Processing, ExplosionAI GmbH, Berlin. Available online: https://spacy.io/ (accessed 21 March 2019).

Fama, E. F. (1980). Agency Problems and the Theory of the Firm. *Journal of Political Economy*, 88(2), 288–307.

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

Farrell, K. A. and Hersch, P. L. (2005). Additions to corporate boards: the effect of gender. *Journal of Corporate Finance*, 11(1-2), 85–106.

FASB (1978). Statement of Financial Accounting Concepts No. 1: Objectives of Financial Reporting by Business Enterprises, Financial Accounting Foundation, Norwalk.

FASB (2010). Conceptual Framework for Financial Reporting: Chapter 1, The Objective of General Purpose Financial Reporting, and Chapter 3, Qualitative Characteristics of Useful Financial Information, Financial Accounting Foundation, Norwalk.

FASB (2019). FASB Taxonomy, Financial Accounting Foundation, Norwalk. Available online: https://www.fasb.org/cs/ContentServer?c=FASBContent_C&cid=1176171809204&d=&pagename=FASB%2FFASBContent_C%2FGeneralContentDisplay (accessed 19 March 2019).

Fenwick & West LLP (2016). Comment Letter - Concept Release: Business and Financial Disclosure Required by Regulation S-K, Fenwick & West LLP. Available online: https://www.sec.gov/comments/s7-06-16/s70616-349.pdf (accessed 19 March 2019).

Filzen, J. J. and Peterson, K. (2015). Financial Statement Complexity and Meeting Analysts' Expectations. *Contemporary Accounting Research*, 32(4), 1560–1594.

Financial Reporting Council (2018). The UK Corporate Governance Code, Financial Reporting Council, London. Available online: https://www.frc.org.uk/getattachment/88bd8c45-50ea-4841-95b0-d2f4f48069a2/2018-UK-Corporate-Governance-Code-FINAL.pdf (accessed 19 March 2019).

Flood, J.M. (2018). Wiley GAAP 2018: Interpretation and application of generally accepted accounting principles, John Wiley & Sons Inc, Hoboken, New Jersey.

Fondas, N. (2000), *Women on Boards of Directors: Gender Bias or Power Threat?*, in Burke, R.J. and Mattis, M.C. (Eds.), Women on Corporate Boards of Directors: International Challenges and Opportunities, Issues in Business Ethics, Springer, Dordrecht, 171–177.

Forbes, D. P. and Milliken, F. J. (1999). Cognition and Corporate Governance: Understanding Boards of Directors as Strategic Decision-Making Groups. *Academy of Management Review*, 24(3), 489–505.

Frederickson, J. R., Hodge, F. D. and Pratt, J. H. (2006). The Evolution of Stock Option Accounting: Disclosure, Voluntary Recognition, Mandated Recognition and Management Disavowals. *The Accounting Review*, 81(5), 1073–1093.

Fülbier, R. U. and Kuschel, P. (2012). Komplexitätszunahme in der IFRS-Rechnungslegung?: Versuche der Systematisierung und indikatorbasierten Messung. *Der Betrieb*, 65(17), 929–937.

Fülbier, R. U., Rupertus, H. and Bielig, J. (2017). SAPs und CAPs - Wesentliche Bilanzierungs- und Bewertungsmethoden, Ermessensentscheidungen sowie Schätzunsicherheiten nach IFRS und US-Normen. *Internationale und kapitalmarktorientierte Rechnungslegung*, 11, 479–486.

Gabaldon, P., Anca, C. de, Mateos de Cabo, R. and Gimeno, R. (2016). Searching for Women on Boards: An Analysis from the Supply and Demand Perspective. *Corporate Governance: An International Review*, 24(3), 371–385.

Galbreath, J. (2018). Is Board Gender Diversity Linked to Financial Performance? The Mediating Mechanism of CSR. *Business & Society*, 57(5), 863–889.

Ghosh, C., Petrova, M.T., Sun, J. and Xiao, Y. (2016). Increasing Gender Diversity in Corporate Boards: Are Firms Catering to Investor Preferences? *Working Paper*, University of Connecticut and Bocconi University. Available online: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2870077 (accessed 19 March 2019).

Gietzmann, M. B. and Trombetta, M. (2003). Disclosure interactions: Accounting policy choice and voluntary disclosure effects on the cost of raising outside capital. *Accounting and Business Research*, 33(3), 187–205.

Gillan, S. L. and Starks, L. T. (2000). Corporate governance proposals and shareholder activism: the role of institutional investors. *Journal of Financial Economics*, 57(2), 275–305.

Glendening, M. (2017). Critical accounting estimate disclosures and the predictive value of earnings. *Accounting Horizons*, 31(4), 1–12.

Glendening, M., Mauldin, E. and Shaw, K. W. (2019). Determinants and Consequences of Quantitative Critical Accounting Estimate Disclosures. *The Accounting Review*, Online Early.

Glendening, M.R. (2012). Critical accounting estimate disclosures and the value relevance of balance sheet items. *Dissertation*, University of Iowa. Available online: https://ir.uiowa.edu/cgi/viewcontent.cgi?article=3300&context=etd (accessed 19 March 2019).

Glick, P. (1991). Trait-based and sex-based discrimination in occupational prestige, occupational salary, and hiring. *Sex Roles*, 25(5-6), 351–378.

Greenberg, R. R., Johnson, G. L. and Ramesh, K. (1986). Earnings versus Cash Flow as a Predictor of Future Cash Flow Measures. *Journal of Accounting, Auditing & Finance*, 1(4), 266–277.

Grosvold, J. (2011). Where Are All the Women? Institutional Context and the Prevalence of Women on the Corporate Board of Directors. *Business & Society*, 50(3), 531–555.

Grosvold, J. and Brammer, S. (2011). National Institutional Systems as Antecedents of Female Board Representation: An Empirical Study. *Corporate Governance: An International Review*, 19(2), 116–135.

Grosvold, J., Brammer, S. and Rayton, B. (2007). Board diversity in the United Kingdom and Norway: an exploratory analysis. *Business Ethics: A European Review*, 16(4), 344–357.

Grosvold, J., Rayton, B. and Brammer, S. (2016). Women on Corporate Boards. *Business & Society*, 55(8), 1157–1196.

Gul, F. A., Srinidhi, B. and Ng, A. C. (2011). Does board gender diversity improve the informativeness of stock prices? *Journal of Accounting and Economics*, 51(3), 314–338.

Gunning, R. (1952). The Technique of Clear Writing, McGraw-Hill Book Company, Inc., New York.

Gupta, V. K., Turban, D. B., Wasti, S. A. and Sikdar, A. (2009). The Role of Gender Stereotypes in Perceptions of Entrepreneurs and Intentions to Become an Entrepreneur. *Entrepreneurship Theory and Practice*, 33(2), 397–417.

Hambrick, D. C. (2007). Upper Echelons Theory: An Update. *Academy of Management Review*, 32(2), 334–343.

Hambrick, D. C. and Mason, P. A. (1984). Upper Echelons: The Organization as a Reflection of Its Top Managers. *The Academy of Management Review*, 9(2), 193–206.

Harrigan, K. R. (1981). Numbers and Positions Of Women Elected to Corporate Boards. *Academy of Management Journal*, 24(3), 619–625.

Hartford Financial Service Group Inc. (2009). Annual Report (10-K), Hartford Financial Service Group Inc. Available online: https://www.sec.gov/Archives/edgar/data/874766/000095012310015756/c96333e10vk.htm (accessed 19 March 2019).

Healy, P. M. and Palepu, K. G. (1993). The effect of financial disclosure strategies on stock prices. *Accounting Horizons*, 7(1), 1–11.

Healy, P. M. and Palepu, K. G. (2001). Information asymmetry, corporate disclosure, and the capital markets: A review of the empirical disclosure literature. *Journal of Accounting and Economics*, 31(1-3), 405–440.

Healy, P. M. and Palepu, K. G. (2003). The Fall of Enron. *Journal of Economic Perspectives*, 17(2), 3–26.

Heckman, J. J., Ichimura, H. and Todd, P. E. (1997). Matching As An Econometric Evaluation Estimator: Evidence from Evaluating a Job Training Program. *The Review of Economic Studies*, 64(4), 605–654.

Hellman, N., Carenys, J. and Moya Gutierrez, S. (2018). Introducing More IFRS Principles of Disclosure – Will the Poor Disclosers Improve? *Accounting in Europe*, 15(2), 242–321.

Henry, T. F. and Holtzmann, M. P. (2006). Critical Accounting Policy Disclosures for Financial Institutions: A Review of Current Practice, Culled from 69 Financial Institutions' 2004 Disclosures. *Bank Accounting & Finance*, 19(3), 14–47.

Herdman, R.K. (2002). Speech by SEC Staff: Critical Accounting and Critical Disclosures, Securities and Exchange Commission. Available online: https://www.sec.gov/news/speech/spch537.htm (accessed 19 March 2019).

Hering, J. (2016). The Annual Report Algorithm: Retrieval of Financial Statements and Extraction of Textual Information. *Working Paper*, University of Erlangen-Nürnberg. Available online: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2870309 (accessed 19 March 2019).

Hering, J. (2018). Textual Analysis in Accounting and Finance, Dissertation, Nürnberg.

Hermalin, B. and Weisbach, M. (1988). The Determinants of Board Composition. *RAND Journal of Economics*, 19(4), 589–606.

Hermalin, B. and Weisbach, M. (2003). Board of directors as an endogenously determined institution: A survey of the economic literature. *Economic Policy Review*, 9(1), 7–26.

Herrmann, D. and Thomas, W. B. (2000). An Analysis of Segment Disclosures under SFAS No. 131 and SFAS No. 14. *Accounting Horizons*, 14(3), 287–302.

Hillman, A. J., Shropshire, C. and Cannella, A. A., JR. (2007). Organizational Predictors of Women on Corporate Boards. *Academy of Management Journal*, 50(4), 941–952.

Hoberg, G. and Phillips, G. (2010). Product Market Synergies and Competition in Mergers and Acquisitions: A Text-Based Analysis. *Review of Financial Studies*, 23(10), 3773–3811.

Holton, V. M. (2000), *Taking a Seat on the Board: Women Directors in Britain*, in Burke, R.J. and Mattis, M.C. (Eds.), Women on Corporate Boards of Directors: International Challenges and Opportunities, Issues in Business Ethics, Springer, Dordrecht, 145–155.

Holtzmann, M. P. (2007). Reporting Critical Accounting Policies. *The CPA Journal*, 42–47.

Hope, O.-K. (2003a). Accounting policy disclosures and analysts' forecasts. *Contemporary Accounting Research*, 20(2), 295–321.

Hope, O.-K. (2003b). Disclosure practices, enforcement of accounting standards, and analysts' forecast accuracy: An international study. *Journal of Accounting Research*, 41(2), 235–272.

Hope, O.-K., Hu, D. and Lu, H. (2016). The benefits of specific risk-factor disclosures. *Review of Accounting Studies*, 21(4), 1005–1045.

Hou, K., van Dijk, M. A. and Zhang, Y. (2012). The implied cost of capital: A new approach. *Journal of Accounting and Economics*, 53(3), 504–526.

House, R.J. (2004). Culture, leadership, and organizations: The GLOBE study of 62 societies, SAGE Publications Inc., Thousand Oaks, California.

Huang, J. and Kisgen, D. J. (2013). Gender and corporate finance: Are male executives overconfident relative to female executives? *Journal of Financial Economics*, 108(3), 822–839.

Hughes, S. B., Sander, J. F. and Snyder, J. K. (2009). Critical accounting policy and estimate disclosures: Company response to the evolving SEC guidance. *Research in Accounting Regulation*, 21(1), 19–33.

IASB (2018). The Conceptual Framework for Financial Reporting, IFRS Foundation, London.

IASB (2019). Disclosure Initiative - Accounting Policies, IFRS Foundation, London. Available online: https://www.ifrs.org/projects/work-plan/accounting-policies/ (accessed 19 March 2019).

Inglehart, R. and Norris, P. (2003). Rising tide: Gender equality and cultural change around the world, Cambridge University Press, Cambridge.

Jensen, M. C. (2001). Value Maximization, Stakeholder Theory, and the Corporate Objective Function. *Journal of Applied Corporate Finance*, 14(3), 8–21.

Jensen, M. C. and Meckling, W. H. (1976). Theory of the firm: Managerial behavior, agency costs and ownership structure. *Journal of Financial Economics*, 3(4), 305–360.

Jianakoplos, N. A. and Bernasek, A. (1998). Are Women more risk averse? *Economic Inquiry*, 36(4), 620–630.

Kang, E., Ding, D. K. and Charoenwong, C. (2010). Investor reaction to women directors. *Journal of Business Research*, 63(8), 888–894.

Kilgour, M. A. (2013). The Global Compact and Gender Inequality. *Business & Society*, 52(1), 105–134.

Klein, M. and Fülbier, R. U. (2018). Inside the Black Box of IASB Standard Setting: Evidence from Board Meeting Audio Playbacks on the Amendment of IAS 19 (2011). *Accounting in Europe*, 74(3), 1–43.

Klerman, D. M., Mahoney, P. G., Spamann, H. and Weinstein, M. I. (2011). Legal Origin or Colonial History? *Journal of Legal Analysis*, 3(2), 379–409.

Krippendorff, K. (1978), *Introduction*, in Gerbner, G. (Ed.), The Analysis of Communication Content: Developments in Scientific Theories and Computer Techniques, R. E. Krieger, Huntington, New York, 3–16.

Krippendorff, K. (2019). Content analysis: An introduction to its methodology, 4th ed., SAGE Publications Inc., Thousand Oaks, California.

Krishnan, H. A. and Park, D. (2005). A few good women—on top management teams. *Journal of Business Research*, 58(12), 1712–1720.

Kuschel, P. (2015). Komplexitätszunahme in der Finanzberichterstattung nach IFRS?: Eine theoretische und empirische Analyse, *Betriebswirtschaftliche Forschung im Rechnungswesen*, 1st ed., Peter Lang GmbH Internationaler Verlag der Wissenschaften, Frankfurt am Main.

La Porta, R., Lopez-de-Silanes, F., Shleifer, A. and Vishny, R. W. (1998). Law and Finance. *Journal of Political Economy*, 106(6), 1113–1155.

Lai, K. M. Y., Srinidhi, B., Gul, F. A. and Tsui, J. S. L. (2017). Board Gender Diversity, Auditor Fees, and Auditor Choice. *Contemporary Accounting Research*, 34(3), 1681–1714.

Lambert, R. A., Leuz, C. and Verrecchia, R. E. (2012). Information Asymmetry, Information Precision, and the Cost of Capital. *Review of Finance*, 16(1), 1–29.

Lang, M. and Stice-Lawrence, L. (2015). Textual analysis and international financial reporting: Large sample evidence. *Journal of Accounting and Economics*, 60(2-3), 110–135.

Lau, D. C. and Murnighan, J. K. (1998). Demographic Diversity and Faultlines: The Compositional Dynamics of Organizational Groups. *Academy of Management Review*, 23(2), 325–340.

Lawrence, A. (2013). Individual investors and financial disclosure. *Journal of Accounting and Economics*, 56(1), 130–147.

Lee, P. M. and James, E. H. (2007). She'-e-os: gender effects and investor reactions to the announcements of top executive appointments. *Strategic Management Journal*, 28(3), 227–241.

Lev, B., Li, S. and Sougiannis, T. (2010). The usefulness of accounting estimates for predicting cash flows and earnings. *Review of Accounting Studies*, 15(4), 779–807.

Levine, C. B. and Smith, M. J. (2011). Critical accounting policy disclosures. *Journal of Accounting, Auditing & Finance*, 26(1), 39–76.

Li, F. (2008). Annual report readability, current earnings, and earnings persistence. *Journal of Accounting and Economics*, 45(2-3), 221–247.

Loughran, T. and McDonald, B. (2011). When Is a liability not a liability? Textual analysis, dictionaries, and 10-Ks. *The Journal of Finance*, 66(1), 35–65.

Loughran, T. and McDonald, B. (2014). Measuring Readability in Financial Disclosures. *The Journal of Finance*, 69(4), 1643–1671.

Loughran, T. and McDonald, B. (2016). Textual Analysis in Accounting and Finance: A Survey. *Journal of Accounting Research*, 54(4), 1187–1230.

Loughran, T. and McDonald, B. (2019a). Stage One 10-X Parse Data. Available online: https://sraf.nd.edu/data/stage-one-10-x-parse-data/ (accessed 19 March 2019).

Loughran, T. and McDonald, B. (2019b). Stop Words. Available online: https://sraf.nd.edu/textual-analysis/resources/#StopWords (accessed 19 March 2019).

Loy, T. R. and Rupertus, H. (2018a). Institutional correlates with female board representation. *Finance Research Letters*, 24, 238–246.

Loy, T.R. and Rupertus, H. (2018b). Board Gender Diversity and its Effects on Capital Markets: New International Evidence. *Working Paper*, Bayreuth. Available online: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3128914 (accessed 19 March 2019).

Lu, C.-W., Chen, T.-K. and Liao, H.-H. (2010). Information uncertainty, information asymmetry and corporate bond yield spreads. *Journal of Banking & Finance,* 34(9), 2265–2279.

Maassen, G.F. (1999). An international comparison of corporate governance models: A study on the formal independence and convergence of one-tier and two-tier corporate boards of directors in the United States of America, the United Kingdom and the Netherlands, *Dissertation*, Rotterdam.

Majors, T. M. (2016). The interaction of communicating measurement uncertainty and the dark triad on managers' reporting decisions. *The Accounting Review*, 91(3), 973–992.

Mallin, C.A. (2016). Corporate Governance, 5th ed., Oxford University Press, Oxford.

Marquardt, C. A. and Wiedman, C. I. (2007). Disclosure, Incentives, and Contingently Convertible Securities. *Accounting Horizons*, 21(3), 281–294.

Mayorga, D. M. and Sidhu, B. K. (2012). Corporate disclosure of the major sources of estimation uncertainties. *Australian Accounting Review*, 22(1), 25–39.

Mayring, P. (2010). Qualitative Inhaltsanalyse: Grundlagen und Techniken, 11th ed., Beltz Verlagsgruppe, Weinheim.

McGregor, J. (2014), *An Index Fund that Bets on Women*. *The Washington Post,* 4 June. Available online: https://www.washingtonpost.com/news/on-leadership/wp/2014/06/04/an-index-fund-that-bets-on-women/?noredirect=on&utm_term=.16ba6d82aabd (accessed 19 March 2019).

Merchant, K. A. (2010). Paradigms in accounting research: A view from North America. *Management Accounting Research,* 21(2), 116–120.

Merck & Co. Inc. (2016). Annual Report (10-K), Merck & Co. Inc. Available online: https://www.sec.gov/Archives/edgar/data/310158/000031015817000010/mrk1231201610k.htm (accessed 19 March 2019).

Miller, B. P. (2010). The Effects of Reporting Complexity on Small and Large Investor Trading. *The Accounting Review*, 85(6), 2107–2143.

Miller, C. C., Burke, L. M. and Glick, W. H. (1998). Cognitive diversity among upper-echelon executives: implications for strategic decision processes. *Strategic Management Journal*, 19(1), 39–58.

Mondelez International, I. (2016). Annual Report (10-K), Mondelez International, I. Available online: https://www.sec.gov/Archives/edgar/data/1103982/000119312517055858/d288385d10k.htm (accessed 19 March 2019).

Monks, R.A.G. and Minow, N. (2011). Corporate Governance, 5th ed., Blackwell Business, Cambridge, Massachusetts.

Nemeth, C. J. (1986). Differential contributions of majority and minority influence. *Psychological Review*, 93(1), 23–32.

Neuendorf, K.A. (2017). The Content Analysis Guidebook, 2nd ed., SAGE Publications Inc., Thousand Oaks, California.

Nielsen, S. and Huse, M. (2010a). The Contribution of Women on Boards of Directors: Going beyond the Surface. *Corporate Governance: An International Review*, 18(2), 136–148.

Nielsen, S. and Huse, M. (2010b). Women directors' contribution to board decision-making and strategic involvement: The role of equality perception. *European Management Review*, 7(1), 16–29.

Ojala, H., Paananen, M. and Renders, A. (2011). The Impact of Accounting Regulation Complexity on Financial Reporting Lead-Time. *Working Paper*, Aalto University School of Economics, Brunel University and Maastricht University. Available online: https://www.brunel.ac.uk/__data/assets/pdf_file/0003/145155/aarcTheImpactOfAccountingComplexity.pdf (accessed 19 March 2019).

Oler, D. K., Oler, M. J. and Skousen, C. J. (2010). Characterizing Accounting Research. *Accounting Horizons*, 24(4), 635–670.

Olson, E. (2019), *Slow Gains for Women and Minorities on Boards of Big U.S. Firms, Study Says*. *New York Times*, 17 January. Available online: https://www.nytimes.com/2019/01/15/business/women-minorities-corporate-boards.html (accessed 19 March 2019).

Organisation for Economic Co-operation and Development (2015). G20/OECD Principles of Corporate Governance, Paris. Available online: https://www.oecd-ilibrary.org/governance/g20-oecd-principles-of-corporate-governance-2015_9789264236882-en (accessed 19 March 2019).

O'Shaughnessy, J. and Rasthy, J. (2005). Critical Accounting Policy Disclosures of Selected NASDAQ Companies. *Journal of Accounting and Finance Research*, 13(2), 19–40.

Palepu, K. (1985). Diversification strategy, profit performance and the entropy measure. *Strategic Management Journal*, 6(3), 239–255.

Palepu, K.G., Healy, P.M. and Peek, E. (2016). Business analysis and valuation: IFRS edition, 4th ed., Cengage Learning EMEA, Andover, Hampshire, United Kingdom.

Pan, Y., Wang, T. Y. and Weisbach, M. S. (2015). Learning About CEO Ability and Stock Return Volatility. *The Review of Financial Studies*, 28(6), 1623–1666.

Pan, Y., Wang, T. Y. and Weisbach, M. S. (2016). CEO Investment Cycles. *The Review of Financial Studies*, 29(11), 2955–2999.

Pande, R. and Ford, D. (2011). Gender Quotas and Female Leadership: World Development Report 2012: Gender Equality and Development, World Bank Group. Available online: https://openknowledge.worldbank.org/bitstream/handle/10986/9120/WDR2012-0008.pdf?sequence=1&isAllowed=y (accessed 19 March 2019).

Paprocki, C. and Stone, M.S. (2004). Is the Quality of Critical Accounting Policy Disclosures lower for companies with high information asymmetry? *Working Paper*, University of Alabama. Available online: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=594202 (accessed 19 March 2019).

Peterson, K., Schmardebeck, R. and Wilks, T. J. (2015). The Earnings Quality and Information Processing Effects of Accounting Consistency. *The Accounting Review*, 90(6), 2483–2514.

Pfeffer, J. (1972). Size and Composition of Corporate Boards of Directors: The Organization and its Environment. *Administrative Science Quarterly*, 17(2), 218–228.

Pfizer Inc. (2016). Annual Report (10-K), Pfizer Inc. Available online: https://www.sec.gov/Archives/edgar/data/78003/000007800317000014/pfe-12312016x10kexhibit13.htm (accessed 19 March 2019).

Pitt, H.L. (2002). Public Statement by SEC Chairman: Remarks at the 29th Annual Securities Regulation Institute, Securities and Exchange Commission. Available online: https://www.sec.gov/news/speech/spch536.htm (accessed 19 March 2019).

Pletzer, J. L., Nikolova, R., Kedzior, K. K. and Voelpel, S. C. (2015). Does Gender Matter? Female Representation on Corporate Boards and Firm Financial Performance--A Meta-Analysis. *PloS one*, 10(6), 1-20.

Poon, S.-H. and Granger, C. W. J. (2003). Forecasting volatility in financial markets: A review. *Journal of Economic Literature*, 41(2), 478–539.

Post, C. and Byron, K. (2015). Women on Boards and Firm Financial Performance: A Meta-Analysis. *Academy of Management Journal*, 58(5), 1546–1571.

Powell, G. N. and Butterfield, D. A. (2002). Exploring the influence of decision makers' race and gender on actual promotions to top management. *Personnel Psychology*, 55(2), 397–428.

Progressive Corporation (2009). Annual Report (10-K) Exhibit 13, Progressive Corporation. Available online: https://www.sec.gov/Archives/edgar/data/80661/000119312510044296/dex13.htm (accessed 19 March 2019).

PwC (2016). File Number S7-06-16. Available online: https://www.sec.gov/comments/s7-06-16/s70616-258.pdf (accessed 19 March 2019).

Radjavi, M. (2012), *Quotas for Securing Gender Justice*, in Gröschl, S. and Takagi, J. (Eds.), Diversity Quotas, Diverse Perspectives: The Case of Gender, Routledge, Farnham, Surrey, 119–133.

Regierungskommission Deutscher Corporate Governance Kodex (2017). German Corporate Governance Code, Regierungskommission Deutscher Corporate Governance Kodex, Frankfurt am Main. Available online: https://www.dcgk.de//files/dcgk/usercontent/en/download/code/170214_Code.pdf (accessed 19 March 2019).

Rhode, D. L. and Packel, A. K. (2014). Diversity on corporate boards: How much difference does difference make? *Delaware Journal of Corporate Law*, 39(2), 377–426.

Richardson, S. A., Sloan, R. G., Soliman, M. T. and Tuna, I. (2005). Accrual reliability, earnings persistence and stock prices. *Journal of Accounting and Economics*, 39(3), 437–485.

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

Rose, C. (2007). Does female board representation influence firm performance? The Danish evidence. *Corporate Governance*, 15(2), 404–413.

Rosenbaum, P. R. and Rubin, D. B. (1983). The central role of the propensity score in observational studies for causal effects. *Biometrika*, 70(1), 41–55.

Roulstone, D. T. (1999). Effect of SEC Financial Reporting Release No. 48 on Derivative and Market Risk Disclosures. *Accounting Horizons*, 13(4), 343–363.

Rupertus, H. (2019). Regulation, Reporting Characteristics and Disclosure Compliance of Critical Accounting Policies. *Working Paper*, University of Bayreuth.

Rupertus, H. and Bravidor, M. (2019). Critical Accounting Policy Disclosures and the Identification of Measurement Uncertainties. *Working Paper*, University of Bayreuth and Heinrich Heine University Düsseldorf.

Salton, G., Wong, A. and Yang, C. S. (1975). A vector space model for automatic indexing. *Communications of the ACM*, 18(11), 613–620.

Schein, V. E. and Mueller, R. (1992). Sex role stereotyping and requisite management characteristics: A cross cultural look. *Journal of Organizational Behavior*, 13(5), 439–447.

Schick, A. G., Gordon, L. A. and Haka, S. (1990). Information overload: A temporal approach. *Accounting, Organizations and Society*, 15(3), 199–220.

Schmid, T. and Urban, D. (2016). Women on Corporate Boards: Good or Bad? *Working Paper*, University of Hong Kong and Technical University of Munich. Available online: https://editorialexpress.com/cgi-bin/conference/download.cgi?db_name=AFA2016&paper_id=570 (accessed 19 March 2019).

Schnatterly, K. and Johnson, S. G. (2014). Independent boards and the institutional investors that prefer them: Drivers of institutional investor heterogeneity in governance preferences. *Strategic Management Journal*, 35(10), 1552–1563.

SEC (1998). A Plain English Handbook - How to create clear SEC disclosure documents, Securities and Exchange Commission, Washington D.C. Available online: https://www.sec.gov/pdf/handbook.pdf (accessed 19 March 2019).

SEC (2000). SEC Concept Release: International Accounting Standards, Securities and Exchange Commission, Washington D.C. Available online: http://sec.gov/rules/concept/34-42430.htm (accessed 19 March 2019).

SEC (2001). Cautionary Advice Regarding Disclosure About Critical Accounting Policies, Release Nos. 33-8040; 34-45149; FR-60, Securities and Exchange Commission, Washington D.C. Available online: https://www.sec.gov/rules/other/33-8040.htm (accessed 19 March 2019).

SEC (2002a). Proposed Rule: Disclosure in Management's Discussion and Analysis about the Application of Critical Accounting Policies, Release Nos. 33-8098; 34-45907; International Series Release No. 1258; File No. S7-16-02, Securities and Exchange Commission, Washington D.C. Available online: https://www.sec.gov/rules/proposed/33-8098.htm (accessed 19 March 2019).

SEC (2002b). SEC Issues Statement on Disclosure Requirements for Public Companies, Securities and Exchange Commission, Washington D.C. Available online: https://www.sec.gov/news/headlines/discreqsjan22.htm (accessed 19 March 2019).

SEC (2003a). Interpretation: Commission Guidance Regarding Management's Discussion and Analysis of Financial Condition and Results of Operations, Release Nos. 33--8350; 34-48960; FR-72, Securities and Exchange Commission, Washington D.C. Available online: https://www.sec.gov/rules/interp/33-8350.htm (accessed 19 March 2019).

SEC (2003b). Summary by the Division of Corporation Finance of Significant Issues Addressed in the Review of the Periodic Reports of the Fortune 500 Companies, Securities and Exchange Commission, Washington D.C. Available online: https://www.sec.gov/divisions/corpfin/fortune500rep.htm (accessed 19 March 2019).

SEC (2007a). Commission Guidance Regarding Management's Report on Internal Control Over Financial Reporting Under Section 13(a) or 15(d) of the Securities Exchange Act of 1934, Release Nos. 33-8810; 34-55929; FR-77; File No. S7-24-06, Securities and Exchange Commission, Washington D.C. Available online: https://www.sec.gov/rules/interp/2007/33-8810.pdf (accessed 19 March 2019).

SEC (2007b). Speech by SEC Staff: Feedback from Individual Investors on Disclosure, Securities and Exchange Commission, Washington D.C. Available online: https://www.sec.gov/news/speech/2007/spch011907ljs.htm (accessed 19 March 2019).

SEC (2009). Final Rule: Proxy Disclosure Enhancements, S7-13-09, Securities and Exchange Commission, Washington D.C. Available online: https://www.sec.gov/rules/final/2009/33-9089.pdf (accessed 19 March 2019).

SEC (2011a). Measurement Uncertainty in Financial Reporting: How Much to Recognize and How Best to Communicate It, Securities and Exchange Commission, Washington D.C. Available online: https://www.sec.gov/about/offices/oca/ocafrseries-briefing-measurement.htm (accessed 19 March 2019).

SEC (2011b). Transcript of the Inaugural Public Roundtable for Financial Reporting Series on November 8, 2011, Securities and Exchange Commission, Washington D.C. Available online: https://www.sec.gov/about/offices/oca/frs/ocafrseries110811-transcript.txt (accessed 19 March 2019).

SEC (2016). Concept Release: Business and Financial Disclosure Required by Regulation S-K, Release No. 33-10064; 34-77599; File No. S7-06-16, Securities and Exchange

Commission, Washington D.C. Available online: https://www.sec.gov/rules/concept/2016/33-10064.pdf (accessed 19 March 2019).

Seierstad, C., Warner-Søderholm, G., Torchia, M. and Huse, M. (2017). Increasing the Number of Women on Boards: The Role of Actors and Processes. *Journal of Business Ethics*, 141(2), 289–315.

Shleifer, A. and Vishny, R. W. (1997). A Survey of Corporate Governance. *The Journal of Finance*, 52(2), 737–783.

Shrader, C. B., Blackburn, V. and Iles, P. (1997). Women In Management And Firm Financial Performance: An Exploratory Study. *Journal of Managerial Issues*, 9(3), 355–372.

Singh, V., Point, S., Moulin, Y. and Davila, A. (2015). Legitimacy profiles of women directors on top French company boards. *Journal of Management Development*, 34(7), 803–820.

Singh, V., Vinnicombe, S. and Johnson, P. (2001). Women Directors on Top UK Boards. *Corporate Governance: An International Review*, 9(3), 206–216.

Sloan, R. G. (1996). Do stock prices fully reflect information in accruals and cash flows about future earnings? *The Accounting Review*, 71(3), 289–315.

Smith, M. (2017). Research Methods in Accounting, 4th ed., SAGE Publications Inc., Thousand Oaks, California.

Soto, M. (2009). System GMM estimation with a small sample. *Barcelona Economics Working Paper Series - Working Paper No. 395*, Institut d'Anàlisi Econòmica. Available online: http://www.iae.csic.es/investigatorsMaterial/a111411124703archivoPdf29632.pdf (accessed 19 March 2019).

Sullivan and Cromwell (2002). Comment Letter - Proposed Rule: Disclosure in Management's Discussion and Analysis about the Application of Critical Accounting Policies. Available online: https://www.sec.gov/rules/proposed/s71602/sullivancromwell.htm (accessed 19 March 2019).

Tajfel, H. (1979). Individuals and groups in social psychology. *British Journal of Social and Clinical Psychology*, 18(2), 183–190.

Terjesen, S., Aguilera, R. V. and Lorenz, R. (2015). Legislating a Woman's Seat on the Board: Institutional Factors Driving Gender Quotas for Boards of Directors. *Journal of Business Ethics*, 128(2), 233–251.

Terjesen, S., Couto, E. B. and Francisco, P. M. (2016). Does the presence of independent and female directors impact firm performance? A multi-country study of board diversity. *Journal of Management & Governance*, 20(3), 447–483.

Terjesen, S., Sealy, R. and Singh, V. (2009). Women Directors on Corporate Boards: A Review and Research Agenda. *Corporate Governance: An International Review*, 17(3), 320–337.

Terjesen, S. and Singh, V. (2008). Female Presence on Corporate Boards: A Multi-Country Study of Environmental Context. *Journal of Business Ethics*, 83(1), 55–63.

The Stanford Natural Language Processing Group (2019). Stanford Named Entity Recognizer (NER), The Stanford Natural Language Processing Group. Available online: https://nlp.stanford.edu/software/CRF-NER.html (accessed 21 March 2019).

Tricker, R.I. (2015). Corporate Governance: Principles, Policies and Practices, 3rd ed., Oxford University Press, Oxford.

U.S. Government (2002). Public Law 107–204: Sarbanes-Oxley Act of 2002, U.S. Government, Washington D.C. Available online: https://www.govinfo.gov/content/pkg/PLAW-107publ204/pdf/PLAW-107publ204.pdf (accessed 19 March 2019).

van Knippenberg, D., Dreu, C. K. W. de and Homan, A. C. (2004). Work group diversity and group performance: An integrative model and research agenda. *The Journal of Applied Psychology*, 89(6), 1008–1022.

Warner Media LLC (2009). Annual Report (10-K). Available online: https://www.sec.gov/Archives/edgar/data/1105705/000095012310014479/g21813e10vk.htm (accessed 19 March 2019).

WEF (2016). The Industry Gender Gap: Women and Work in the Fourth Industrial Revolution, World Economic Forum. Available online: http://www3.weforum.org/docs/WEF_FOJ_Executive_Summary_GenderGap.pdf (accessed 19 March 2019).

WEF (2017). The Global Gender Gap Report 2017, World Economic Forum. Available online: http://www3.weforum.org/docs/WEF_GGGR_2017.pdf (accessed 19 March 2019).

Williamson, O. (1984). Corporate Governance. *The Yale Law Journal*, 93(7), 1197–1230.

Wintoki, M. B., Linck, J. S. and Netter, J. M. (2012). Endogeneity and the dynamics of internal corporate governance. *Journal of Financial Economics*, 105(3), 581–606.

Withers, M. C., Hillman, A. J. and Cannella, A. A. (2012). A Multidisciplinary Review of the Director Selection Literature. *Journal of Management*, 38(1), 243–277.

Wolk, H.I., Dodd, J.L. and Rozycki, J.J. (2017). Accounting theory: Conceptual issues in a political and economic environment, 9th ed., SAGE Publications Inc., Thousand Oaks, California.

Wooldridge, J.M. (2013). Introductory Econometrics: A Modern Approach, 5th ed., South-Western Cengage Learning, Mason, Ohio.

Xie, H. (2001). The mispricing of abnormal accruals. *The Accounting Review*, 76(3), 357–373.

Yermack, D. (2006). Board Members and Company Value. *Financial Markets and Portfolio Management*, 20(1), 33–47.

Yeung, P. E. (2009). Uncertainty and expectation revisions after earnings announcements. *Contemporary Accounting Research*, 26(1), 273–301.

Zahra, S. A. and Pearce, J. A. (1989). Boards of Directors and Corporate Financial Performance: A Review and Integrative Model. *Journal of Management*, 15(2), 291–334.