The Coherence Premium

The effect of strategic coherence on financial firm performance

Master's Thesis A.Q. van Braam Houckgeest

Graduation Committee

	Delft University of Technology
Chairman:	Prof.dr. C.P. van Beers
First supervisor:	Dr. E. den Hartigh
Second Supervisor:	Dr.ir. Z. Roosenboom-Kwee
	Booz & Company
First External Supervisor:	Dr. ir. M.R.B. Reunis
Second External Supervisor:	Prof. dr. K. Cools RA

Executive Summary

For many decades the explanation of firm success and failure has been a subject of interest for many scholars. A wide variety of theories have been developed that provide partial explanations of firm performance. But the majority of the variance cannot be explained yet. This research project aims to explain firm success from a different point of view: coherence. The notion of coherence combines elements of existing strategic theories into an overarching coherence theory for success.
The goal of this project is to contribute to knowledge on the coherence premium relationship by performing a large-scale quantitative study, using textual analysis. And indeed evidence for this relationship was found.

Large-scale (automated) quantitative measurement of coherence was not feasible; therefore an approximation was developed using textual analysis. Starting from the coherence theory of Leinwand and Mainardi (2010), coherence was defined as alignment of a firm's core capabilities system, its way to play and its product and service portfolio. Subsequently literature research was performed to expand the coherence theory model, to find other proxies for coherence (elements) and to find other cues for the positive relationship between coherence (elements) and performance. An extended coherence model was developed and operationalized for use with textual analysis. A coherence score could then be calculated by multiplying the core capabilities system, way to play and product and service portfolio elemental scores with each other. The elemental score measured the emphasis on one of the three coherence elements with textual analysis by counting predefined keywords indicative for that coherence element. The analysis was performed with the U.S. Securities Exchange Commission (SEC) 10-K filing item 7 Management Discussion and Analysis (MD&A), in which the firm's management discusses the performance of the past year and future plans, both on a (business unit) strategic level. The scores were taken relative to the length of the MD&A and were benchmarked against a peer group of nearly 7,000 consumer and retail companies. Premium was defined as above normal financial performance compared to peers.

Evidence was found for a positive relationship between coherence of a company and premium with this coherence approximation. A summary of the results for the full large sample is shown in Table 1.
Sales growth is significantly positively related to coherence, the other premium metrics are positively related to coherence as well, but to a limited extent.
The study adds a quantitative basis to the existing coherence related literature on the essence and impact of coherence. The results indicate that more emphasis placed on the coherence elements in the MD&A is positively related to (financial) performance. Results with actual coherence may differ from these findings. Looking at the literature, case studies and coherence profiler survey results, the current results are likely to become stronger, in effect and/or significance, with a more accurate approximation.

Table 1 One-way ANOVA results for the full large sample. The suffixes * and ** indicate that differences between the high and low coherence groups are significant at the 0.10 and 0.01 significance level respectively.

Premium measure	Coherence Group	Mean	Sig.
Sales growth	Low Coherence	9.46%	.007**
	High Coherence	12.09%	
Return on assets	Low Coherence	-0.81%	.087*
	High Coherence	0.29%	
Earnings before interest and taxes margin	Low Coherence	4.46%	.069*
	High Coherence	5.03%	
Price/Earnings ratio	Low Coherence	23.1	.077*
	High Coherence	25.5	

Subsequently the coherence effect was corrected for size, diversification or industry. This was done by testing the premium differences between groups of different industries, between groups of different sizes (by sales) and between diversified and non-diversified firms. It was hypothesized that these relevant factors moderate the relationship between coherence and premium.

- Diversification makes a significant difference. Undiversified firms exhibit stronger relationships for sales growth and P/E ratio. Diversified firms exhibit stronger relationships for ROA and EBIT margin and show insignificant but inversed differences for sales growth and coherence.
- The size groups affect the differences between the coherence groups too. Differences and significance, compared to the full large sample, are enlarged for sales growth and P/E ratio within smaller companies and for ROA and EBIT margin for larger companies. The reverse is true for the other premium measures and the medium size category only deviates from the general results by lower significances. These results clearly differentiate between the effects of coherence in small and large companies.
- Industry (by two-digit SIC-codes) high and low coherence group differences generally resemble the differences found for the full sample. However, there are some distinct exceptions in several SIC-codes. This makes it an important determinant of firm performance, which has to be taken into account

Overall, large and diversified companies perform better than smaller and undiversified companies. The datasets of the large and/or diversified groups consist of a relatively high fraction of highly coherent items. The better performance and the dataset composition could be caused by the survival bias; only the coherent and successful companies grow to become large multinationals, active in many segments.
A related finding is that diversified and larger firm groups mainly consist of high coherent companies, whereas the undiversified and smaller firm groups have more low coherence firms. An explanation can be found in the survival rate. Diversified firms are generally larger than undiversified firms, because it takes time to expand to new markets. Companies need time to grow large, but only the most effective companies, which are (highly) coherent, will survive long enough to reach that size. This implies that coherence and firm survival are positively correlated.

Table of Contents

Preface

This study was performed as a final graduation project for the Master's degree Management of Technology at Delft University of Technology and for Booz & Company. The research project quantitatively tests whether coherence, the final element in an important Booz & Company strategic theory, has a positive relationship with premium, using textual analysis. This study is the first to succeed in finding evidence for the coherence premium theory based on a large quantitative analysis.

From an academic point of view it adds quantitative support for the coherence premium theory and therewith for a large amount of empirical literature focused at coherence related aspects. This research supports the Overarching "Gestalt" School (Venkatraman and Camillus, 1984) and creates a bridge between the Resource Based View and Industrial Organisation theory, because coherence within a firm can be seen as the fit between the capabilities and market characteristics. Stated in coherence terminology the core capabilities system and product and service portfolio should be successfully aligned by implementing the right way to play. Elements of Evolutionary Economics and Dynamic Capabilities schools of thought are implicitly present in the coherence theory as well. This makes it an interesting overarching strategic theory.
In a more practical sense the research performed provides further evidence for the coherence premium as a basis for thinking about strategy within Booz & Company.

The research project opens doors for advanced follow-up research using textual analysis and it supplies new cues that can be used in other coherence related studies.

Looking back on six intensive months of research, the success of the project gradually becomes clearer. First of all the vague concept of coherence was successfully operationalized. At the start of the project, the biggest challenge was thought to be the development of a quantitative measure for coherence. This issue was overcome with large-scale textual analysis. Determination and perseverance was required for the manual coding of large amounts of text, needed for the coherence measure construction. The constructed coherence measure was then pre-tested for a large number of measurement options. This new and original approach was subsequently automated and applied to a large sample of almost 7,000 items. The effective collection, evaluation and cleaning of this dataset provided a powerful basis for the statistical analyses to follow. Finally, the results were very rewarding, because they provided evidence for the hypothesised positive relationship between coherence and premium. For this research project academics and practitioners were united internationally. These connections and the ensuing cooperation have had a major positive influence on the research in hand.

All in all, the past six months felt like a high-speed train with many unexpected and exciting stops, side-tracks and turns. Having safely returned, I am glad to have had the opportunity to join the ride.

Acknowledgements

I am grateful to all who have helped with the successful completion of this thesis: my supervisors, co-workers, reviewers, friends/family and sponsors.

First of all, I would like to thank my supervisors.
Erik den Hartigh, our weekly Skype-meetings were very valuable and enjoyable. I always looked forward to them. You were very engaged and kept me on track throughout the entire project. It was great to have an experienced and down-to-earth academic like you to test my ideas on; during our sessions I really learned how to perform research. Also, I really appreciated your flexibility, even during the weekends.
Zenlin Roosenboom-Kwee, your help with the literature research and especially with the coding was invaluable, as it gave me a clear direction.
Marc Reunis, your questioning has made my thesis much stronger. 'Is this relevant? If so, where do you use it? Why do you take this approach? Why?' Prying questions, to which I often did not have a clear answer. I gradually learned to ask these questions myself and to do this in an earlier stage, for which I am grateful as it has helped me very much. The way you were able to supervise me on top of a job with demanding hours is admirable.
Kees Cools, our sessions were always interesting as you managed to take another point of view, raise new issues and find new opportunities very well.

Special thanks to Joost Impink, who has helped me tremendously by writing the perl scripts which enabled the automation of the data acquisition and coherence score calculation. Our quick email exchanges, about the script as well as about other aspects of the research, were very pleasant.

As the work atmosphere is a key component for success, I would like to thank all my colleagues at Booz & Company for our very enjoyable personal as well as research related encounters. The atmosphere was great and I learned a lot from you all.

Furthermore, I would like to express my gratitude towards my ''readers'' who helped me improve this thesis by providing very accurate and helpful feedback on its contents and presentation.
Last, thanks to my family, friends, roommates and Linda; you all helped a lot as sparring partners and by taking my mind off things.

Finally, this research would not have been possible without the support of Booz & Company. I would especially like to thank them for enabling the international correspondence and cooperation, which brought the project to a higher level. Booz & Company, thank you for your support.

A.Q. (Andreas) van Braam Houckgeest
Delft, the Netherlands
August 30, 2012

List of Figures

List of Tables

1. Introduction

This first chapter introduces the subject by providing background information on the project followed by a discussion of the research objectives. Then the main research question and its sub questions are revealed in Section 0. Subsequently the scope is delineated in Section 1.4. Then in Section 0 an initial research model is discussed. Finally, the general methodology and an outline for this thesis are presented.

1.1 Project Background

For many decades the explanation of firm success and failure has been a subject of interest for many scholars. In this process several schools of thought have been established. Within and in between them a wide variety of theories have been developed that can provide partial explanation of firm performance. But the majority of the variance cannot be explained yet. The goal of this research project is to explain firm success from a different point of view: coherence. The notion of coherence combines elements of existing strategic theories into a new generic framework for success. The coherence idea comes from industry; it has been developed with wide experience in strategy consulting.

'The Coherence Premium' is a recently published article by two senior partners at Booz & Company (Leinwand and Mainardi, 2010). The capabilities-driven strategy and coherence theory are described as the backbone for the Booz & Company view on business strategy. It was derived from many years of consulting practice. Therefore the support for the theory is primarily based on their experience. It is backed up by several case studies performed in different industries. In the article it is stated that *'coherent companies reap sustainable and superior returns from a capabilities-driven strategy'*. This is substantiated by a plot of 'Coherence' against EBIT margin from which a positive correlation between coherence and EBIT margin is concluded.
The effect of coherence - essentially a fit between company strategy, capabilities and environment - on a firm's financial performance is a subject with important managerial implications. However, the assumed causalities, substantiating the 'coherence premium', have yet to be scientifically tested for validity. 'The Coherence Premium' article serves as a starting point for this research project.

Content-wise the article revolves around coherence, which is built up by strategic fit or alignment amongst the three elements way to play, core capabilities system and product and service portfolio. The concepts of strategic fit, strategic alignment and capabilities have been previously investigated by scholars. However, usually the focus is on developing theoretical constructs or in-depth research of a particular field or subject, like information technology, innovation, entrepreneurship or company structure. Typically, the research performed is focused on a single industry or a specific type of firm. The more general concept of a capabilities-driven strategy resulting in a coherence premium can be applied to most, if not all, industries and practices. Therefore it adds value to current strategy research as a new overarching theory in between current strategy schools.

1.2 Research Objectives

The goal of this project is to contribute to knowledge on the coherence premium by performing a large-scale quantitative study. Stated otherwise, *is coherence positively related to the financial performance of a firm?* An additional goal is to assess the potential of textual analysis.

1.3 Research Questions

The primary research question is stated as:

What is the relationship between 'coherence' in a company and the 'premium' it delivers?

In this question two terms, 'coherence' and 'the premium' have to be conceptualised. Moreover, a rationale behind the expected positive relation between coherence and premium must be developed.

In Table 2 the research sub questions are listed, together with their approach to find an answer and their position in the thesis. The concepts of coherence and premium are explained in Chapter 2. In the literature review, Chapter 3, relevant publications are discussed in search for an extension of the coherence concept and for results indicating relationships between coherence and premium. Chapter 4 then uses the information found to develop an operationalized framework. Several measurement options are indicated here. These measurement options are used to develop the methodology for measuring coherence and premium with textual analysis in Chapter 4. The results from the performed analyses are then presented in Chapter 7.

Table 2 Research questions with the approach to find an answer and their position in this thesis.

	Research Question	***Approach***	***Chapter***
1a	What is coherence?	Literature	2, 3
b	How can it be measured?	Literature, Textual analysis	2 - 0
2a	What is premium?	Literature	2, 3
b	How can it be measured?	Literature	3, 4
3	What is the expected relationship between coherence and premium?	Literature	3, 4
4	How can the relationship of Q3 be quantitatively investigated?	Methodology: textual analysis	4
5	What are the relationships found between coherence and premium?	Statistics	7

1.4 Scope

The article 'The Coherence Premium' (Leinwand and Mainardi 2010), served as the fundamental starting point for this research. The textual analysis was confined to 10-K SEC-filings of eight sectors from the following FamaFrench-48 industries: food, beer, retail, meals, soda, clothes, textile and paper. These eight sectors are part of the consumer and retail industry, like the companies in Leinwand and Mainardi (2010). These eight sectors form the large sample *peer group* in this research.

Coherence theory suggests a causal relationship between coherence and premium. The current study enables investigation of the relationship between coherence and premium. It does not aim to investigate and draw conclusions regarding the direction of the causality.

In Chapter 8 the starting point and other initial scope decisions are reviewed to evaluate their appropriateness and their (potential) effects on the results obtained.

1.5 Initial Model

In this section the starting model and the hypothesis will be presented. In the first section of Chapter 4 this model is extended and operationalized.

The starting model for the coherence premium relationship is shown in Figure 1.

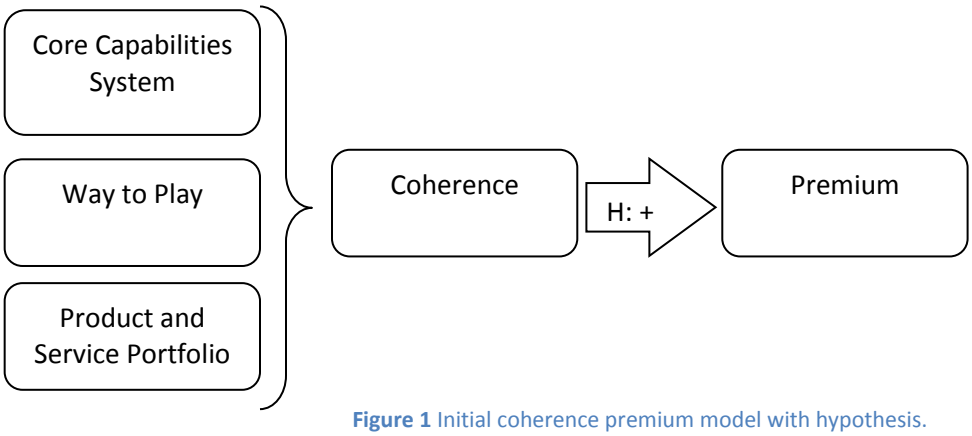

Figure 1 Initial coherence premium model with hypothesis.

The initial hypothesis is that coherence is positively related to premium. Initial argumentation comes from:

- Leinwand and Mainardi (2010), the starting point for our research;
- Booz & Company strategy consulting experience.

The rationale for the positive relationship between coherence and premium will be further extended throughout this thesis.

1.6 General Methodology and Outline

First of all, the Leinwand and Mainardi (2010) notion of coherence, the departure point for this research, is explained. Subsequently it is given a place in strategy literature, after which a general coherence framework is developed. This is Chapter 2.

To improve and extend this framework an extensive literature review was performed in Chapter 3. This is executed using a procedure which maps relevant literature findings onto the general coherence framework.

The research methodology is explained in Chapter 4. With the information obtained from the literature mapping the framework is extended and operationalized in Section 4.1. Subsequently the choice for textual analysis will be explained in Section 4.2. In the last section the analysis method is discussed.

Then, in Chapter 5, a measure to approximate coherence, using textual analysis, is developed. A *test set*[1], consisting of 36 company-year data points used in Leinwand and Mainardi (2010), is used to test the developed coherence measure before applying it to the *large sample*, which consists of almost 7,000 data points from the peer group. This test set helps to assess and choose between the many different options that are available in the design of the coherence measure.

In Chapter 8 the large dataset is analysed, discussed and cleaned; it is prepared for the analyses.

The results of these analyses performed with the large sample are presented and discussed in Chapter 7.

Finally, in Chapter 8 conclusions are drawn and the obtained results are put into perspective. Limitations are discussed and recommendations for further research are given here as well.

[1] More information about the test set can be found in Appendix II.

2. The Coherence Premium

In this chapter the starting point of this research project: 'The Coherence Premium' will be introduced and explained. The coherence premium concept is given a place in the strategy research field, by discussing several relevant meta-studies and frameworks in Section 2.2. Subsequently, the general coherence premium model is derived from the information covered in the earlier sections. It is presented in the last section.

2.1 What is Coherence?

In this section the coherence theory developed by Leinwand and Mainardi (2010, 2011) is explained. In the first section coherence is defined at the highest level. Subsequently, the elements coherence consists of are explained in Section 2.1.2. The third section explains how Leinwand and Mainardi (2010) measured coherence. Why coherence leads to a premium is explained in the last section.

2.1.1 Definition

Leinwand and Mainardi give two definitions for coherence:

- Leinwand and Mainardi (2010, page 88): *'A coherent company focuses on what it does best in making every decision across every business. It derives its 'right to win' from aligning the way to play, the capabilities system and the product & service fit'*.
- The Essential Advantage (2011a, page 22): ***'**A resolute, clear-minded focus in a company on three critical elements: its way to play, its most distinctive capabilities, and its line-up of products and services. The better aligned these three elements, the more coherent the company.'*

Coherence involves three elements: the *way to play*, the *core capabilities system* and the *product and service portfolio*. If these elements are properly combined and fit together this results in coherence. In other words: Coherence is the fit between a *way to play*, a *core capabilities system* and the *products and service portfolio*. Figure 2 depicts the interaction between the three coherence elements.

Figure 2 Visual representation of the coherence elements. Reproduced from Booz & Company.

In the next section the three elements that, together, define coherence are explained.

2.1.2 Coherence Elements

In this section the coherence elements way to play, the core capabilities system and the product and service portfolio are explained.

Way to Play

'A way to play is a considered approach for creating and capturing value in a particular market, in a way that differentiates the company from all other companies. A well-defined way to play is broad enough to allow flexibility and growth, but narrow enough to focus strategy and decision making. It may involve being an innovator, a low-cost provider, a premium player, or (more likely) a distinctive combination of strategic identities that fits your customers and capabilities precisely' (Leinwand and Mainardi, 2011a, page 23).

Thus, a way to play incorporates strategic direction, business models and value propositions. It can be seen as the company's strategic plan. A way to play can be found by combining several so-called strategic *'puretones'*, shown in Table 3, identified by Leinwand and Mainardi (2011a). This can make ways to play very specific. A simplification is to look at single puretones only. However, with fifteen puretones thisis not specific enough to represent a real way to play.

Table 3 The fifteen puretones identiefied by Leinwand and Mainardi. Adapted from Leinwand and Mainardi 2011a, page 71-74.

Puretone	Value Proposition
Aggregator	Provides the convenience of a one-stop solution
Category Leader	Maintains top market share in a category and uses that position to shape and influence downstream channels as well as upstream supply markets
Consolidator	Dominates an industry through acquisitions ("rolling up an industry") to provide either a value benefit to consumers or access to a platform with products and services that otherwise would not be possible
Customizer	Leverages insight and market intelligence to offer tailored products or services
Disintermediator	Helps customers bypass unreachable or more expensive distribution channels and parts of the value chain, thereby providing access to otherwise inaccessible services and products
Experience Provider	Builds enjoyment, engagement, and emotional attachment through strong brands or experiences
Fast Follower	Leverages foundations laid by innovators to quickly introduce competing offerings, often at greater value or to a broader base of consumers
Innovator	Introduces new and creative products or services to the market
Platform Provider	Operates and oversees a shared resource or infrastructure
Premium Player	Offers high-end products or services
Reputation Player	As a trustworthy provider, charges a premium or gains privileged access to customers
Regulation Navigator	Offers access to otherwise unreachable products and services by managing within government rules and oversight, and by influencing them
Risk Absorber	Mitigates or pools market risk for its customers
Solutions Provider	Provides bundled products and services that fully address an unmet customer need
Value Player	Offers lowest prices or tremendous value for comparable products and services

Core Capabilities System

A Capability is *'Something that customers value, which a company does well and its competitors cannot beat. It is more than an activity or a function; it is the interconnection of people, knowledge, IT, tools, and processes that enable a company to outexecute rivals on some important measure'* (Leinwand and Mainardi, 2010, page 88).

A core capabilities system is *'The engine of value creation; it is the system of three to six capabilities that allow companies to deliver their value proposition'* (Leinwand and Mainardi, 2010, page 88).
'Each of the capabilities within this system is distinctive: it represents an extraordinary competence that few others can master. When deployed together, in a way that is relevant to the corporate strategy, capabilities can enable the company to consistently outperform rivals. It is not just the assets that matter most, not even the products and services sold. It is what you do, time and again, for the customers you serve. A capability is ensured through the right combination of processes, tools, knowledge, skills and organisation all focused on meeting the desired result.' (Leinwand and Mainardi, 2011a, page 23).

A core capability is something a company does best, a best practice, which is valued by the users of the products/services this company delivers. A core capabilities system is several of these core capabilities working together to deliver value for customers.

Leinwand and Mainardi use very specific core capabilities as examples in their work. For Wal-Mart for instance, they identify four core capabilities: aggressive vendor management, expert point-of-sale data analytics, superior logistics and rigorous working capital management. There are potentially hundreds of these kind of capabilities.

Product and Service Portfolio
The product and service portfolio is coherent if *'all products and services leverage the same capabilities system'* (Leinwand and Mainardi, 2010, page 88). It is *'the degree to which the line-up of products and services is supported by a unique capabilities system and aligned with the way to play'* (Leinwand and Mainardi, 2011a, page 23).

2.1.3 Measuring Coherence

Measuring coherence is not straightforward. '*Coherence can be estimated by assessing the extent to which a company's products and services share the same distinctive capabilities.'* (Leinwand and Mainardi, 2011a, page 22).
In 'The Coherence Premium' coherence is scored as follows: '*First the segments each company serves are defined. Next the capabilities that drive value for the company in each segment are defined. Finally the number of common capabilities across all segments the company serves is determined*' (Leinwand and Mainardi, 2010, page 91-92). This is a highly labour intensive method which requires deep insight into the firm under investigation, because the capabilities that drive value in each segment are often hard to reveal. Some capabilities are hidden and sometimes even the company itself is unaware of its value-adding capabilities.

2.1.4 Why Coherence Results in Premium

The coherence premium proposition is that if a company is (relatively) coherent, it will attain a financial premium. In Leinwand and Mainardi (2010) this premium is defined as the Earnings Before Interest and Taxes (EBIT) margin, which equals EBIT divided by net revenue, over a five-year period.

Leinwand and Mainardi (2010, adapted from pages 90-91) identify four ways in which coherence creates value:

- Coherence strengthens a company's competitive advantage. Companies that focus on their capabilities, day in day out, continually improve them. Employees become more skilled and systems grow more adept, enabling companies to consistently outperform their rivals.

- Coherence focuses strategic investment on what matters. Companies make better organic growth decisions and pursue acquisitions that are in line with their capabilities. Coherent companies direct capital, time and talent to those activities, products and businesses that will extend their lead.
- Coherence produces efficiencies of scale. Companies can spend more wisely and grow more easily when they deploy the same capabilities across a larger array of products and services. Furthermore, these companies can apply the capabilities to businesses that would not normally be able to afford them.
- Coherence creates alignment between strategic intent and day-to-day decision making. Coherent companies execute better and faster because everyone in the organisation understands what is important.

After this introduction to the coherence premium the positioning of coherence within the strategy research field will be discussed.

2.2 Positioning of Coherence

Now that the basic concept of coherence has been explained and the hypothesis has been formulated, a helicopter view on strategy research will be adopted to find the positioning of coherence within the strategy research field. General findings from three meta-studies are presented in Section 2.2.2, followed by a general framework describing most of the relationships investigated. Then the concepts of strategic fit and strategic alignment will be introduced, which are closely related to coherence. Finally the place of the coherence premium concept in strategy research will be discussed in Section 2.2.6. Some background in the strategy field is needed, as a refresher the main schools of thought, and the position of coherence within these schools, is given in the first paragraph.

2.2.1 General Strategic Theories

During the past decades many strategic theories have been developed to explain (sustained) firm success and failure. First the four presently dominant strategic views will be presented briefly to provide more background to strategic research. Then another relevant theory for this research, contingency theory, which is often used in strategy research, will be explained.

Four Main Strategic Streams

The four most widely accepted strategy perspectives (Zegveld, 2004) are: Industrial Organisation theory (IO), Evolutionary Economics (EE) the Resource-based view (RBV) and the Dynamic Capabilities framework (DC). The differences between these strategic schools lie in the scope.

- In RBV one starts with the firm internally. After analysing the firm a suitable market to compete in should be found.
- IO theory starts at the other end, looking at the competitive environment and adapting the firm to the environmental implications.
- EE takes the external point of view, but is more focused on the influence of the time component, for instance by looking at path dependencies.
- DC revolves around the evolution (creation, destruction, enhancement etc.) of capabilities within the firm. More specifically, the capability to reinvent, reinforce, develop and build on existing capabilities is regarded essential in a turbulent market environment.

The four strategy perspectives can be arranged in a matrix of two dimensions: point of view (internal or external) and time orientation (contemporary or path dependence). It is shown in Figure 3.

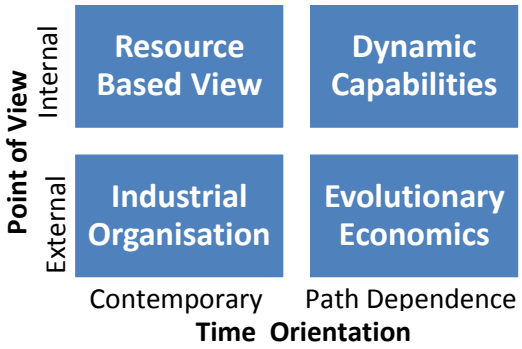

Figure 3 The four main strategy perspectives, positioned by point of view and time axes.

Theories that take an external reference point argue that performance is caused by the firm's industry environment and the firm's position in the marketplace. Theories based on an internal reference point on strategy state that performance is caused by the firm's resources, in its widest definition. Contemporary theories do not explicitly take the past or future into account, whereas path dependence oriented theories stress the importance of the path through which the strategic outcomes developed. They incorporate the effects of change on performance, for instance by path dependencies which explain how current strategy is limited by the decisions made in the past, even though the circumstances may have changed considerably. All four theories will be briefly explained in the next sections.

Industrial Organisation

IO theory has an external point of view and a contemporary time orientation. According to this theory the performance of a firm results from the overall industry attractiveness and its competitive position within that industry. The most famous framework based on IO theory probably is the five competitive forces framework of Michael Porter, which is used to define market attractiveness (Porter 1979). It includes competition amongst existing firms, power of suppliers and customers, and threat of new entrants and substitutes to evaluate the market attractiveness. Furthermore Porter found three general competitive strategies to obtain competitive positions: cost leadership, differentiation and focus (Porter 1980). Perhaps, more importantly, he found that combinations of these strategies generally do not work.

The positioning of coherence theory within IO is limited. Although coherence takes into account the market characteristics in its element product and service portfolio, it does not matter in which industry or in how many industries the firm is active for coherence. However, the internal characteristics do matter, because an effective way to play should combine them with the core capabilities of the company to achieve an optimal match and create maximum value. It can be concluded that in IO the way to play is contingent on the market characteristics.

Resource-Based View

The RBV has an internal point of view and a contemporary time orientation. It explains competitive advantage from the resources of a company, e.g. machinery, people, procedures, capabilities. A core capabilities system can be seen as a resource in this context. Barney (1991) argues that resources in itself are not enough to gain competitive advantage; they must be VRIN: valuable, rare, non-imitable and non-substitutable. Newbert (2007) concluded that the RBV, although widely appreciated, has only received modest support, with a wide variability amongst independent variables and theoretical approaches. In the same year, Chmielewski and Paladino (2007) demonstrated that a resource

orientation resulted in higher ROC and ROS with an online survey amongst 149 consumer good firms. Newbert (2008) researched the relationships between value, rareness, competitive advantage and performance at a conceptual level. He found that value and rareness are related to competitive advantage. Furthermore he found competitive advantage to be related to performance and to mediate the rareness-performance relationship.

Coherence is closely related to RBV by the core capabilities system, which is the foundation of coherence theory. Coherence is therefore resource oriented and the core capabilities system must be VRIN to add most value. Coherence theory goes one step further than RBV, by using a way to play, which should be VRIN as well, to maximise the added value from its resources/core capabilities via the coupling with markets that value these resources/core capabilities. Essentially coherence is about taking advantage of your best practices and applying them to products and services where they add most value. This causes a competitive advantage, which can only be sustained if the combination of capabilities, way to play and products is VRIN. If it is not VRIN, competitors can develop and use the same capabilities and ways to play, eliminating the initial competitive advantage. Coherence theory can thus be seen as an extension of the RBV, which adds an explanation, the way to play, for the step from competitive advantage towards improved (financial) performance.

Evolutionary Economics

EE theory has an external point of view and a path dependence time orientation. Performance is explained by firm actions and responses to events in the marketplace. According to Barnett and Burgelman (1996), EE means developing dynamic, path-dependent models that allow for random variation and selection within and between organizations. Game theory can be incorporated to assess the effect of competitor responses. EE can also be regarded a competitive race in which high performance comes from speed and innovation and this performance must be taken relatively to the industry competitors. This view is also known as ‘red queen competition’, where firm performance can only be increased by outperforming competitors (Derfus et al. 2008).

Coherence is not directly linked to the EE theory, as it does not explicitly take into account path-dependency and tactical decision making like competitor responses. These aspects will have an effect, because they have an implicit influence on all three coherence elements. Therefore coherence in EE theory is about the fit of the coherence elements with past, present and future events, actions and responses. Examples can be found in the payoffs of a certain way to play that can change over time or buying and building a new core capability.

Dynamic Capabilities

The DC framework has an internal point of view and a path dependence time orientation. It is important for industries that change rapidly and unpredictable, because RBV cannot explain competitive advantage in these markets. Eisenhardt and Martin (2000) define dynamic capabilities as the organizational and strategic routines by which companies achieve new resource configurations. It is the ease with which an organisation and its employees can learn, adapt and change. Capabilities can be imitated. Therefore the value lies in the creation of new resource configurations and not in the capabilities themselves. Evolution of dynamic capabilities is often described as being path dependent. This path dependency can be explained by learning mechanisms. Examples include repeated practice, codification of experience and learning from mistakes.

Recently Teece (2007) further developed the DC concept. He states that dynamic capabilities drive sustained competitive advantage by enabling the creation, deployment and protection of the intangible assets vital to superior business performance. He defines the distinct skills, processes, procedures,

organizational structures, decision rules, and disciplines as micro foundations of dynamic capabilities, which result in sensing, seizing and reconfiguring capacities at the enterprise level. Enterprises with strong dynamic capabilities adapt well to the environment, but also shape their environment through innovation and collaboration (Teece, 2007).

The core capabilities system in coherence theory should be reinforced continually, it should be applied to all products and services and continuous learning and adaptation are very important. To achieve coherence can actually be seen as a dynamic capability itself. Dynamically achieving and sustaining or improving the fit between resources/capabilities, environment, strategy and more, taking into account historic, present and future events, is what coherence is about. Therefore the core capabilities element is closely related to DC. Coherence theory thus incorporates DC as well.

Contingency Theory

According to the contingency theory firm performance is contingent on a large number of specific factors. Therefore no single best strategy, applicable to all firms, exists. The optimal strategy is found to be contingent (dependent) upon both the internal and the external situation and environment. Research in this area often links process, context and structure variables to performance indicators, (Hansen and Wernerfelt, 1989; Olson et al., 2005) this is called Structural Contingency Theory. Miller (1992) concludes that there often exists a trade-off between internal fit, the firm's structure, culture and processes, and external fit, alignment between the firm and its surrounding market conditions. This implies that managers have to choose for one of the two.
Drazin and Van de Ven (1985) posed that at least three different conceptual approaches to fit were used:

- The selection approach: Fit is caused by the evolutionary process of adaptation that ensures that only the best-performing organizations survive. It is closely related to EE.
- The interaction approach: Variations in organizational performance are explained with the interaction of organizational structure and context. A deviation-score approach has often been used. Instead of looking for classical interaction effects, the influence of deviations in structure from an ideal context-structure model is analysed. Fit is defined as adherence to a linear relationship between dimensions of context and structure. A lack of fit results from a deviation from this relationship. This approach comes close to IO theory.
- The systems approach: It emphasizes the need to adopt multivariate analysis to examine patterns of consistency among dimensions of organizational context, structure, and performance. It is related to the DC framework, as it involves the continuous realignment among different dimensions.

They have a major impact on the meaning of contingency theory and the resulting empirical findings. Comparing these three methods, support was found for the selection and systems approach, but not for the interaction approach (Drazin and Van de Ven, 1985).

Coherence theory agrees with contingency theory on the notion that no single best, widely applicable, strategy exists. The optimal strategy is contingent on a variety of factors. However, coherence theory does not (explicitly) take into account the structural contingency variables, such as age, size and structure. These are implicitly used; the core capabilities system is dependent on these firm characteristics and the way to play is influenced as well. Coherence in the light of contingency theory can be seen as the fit between strategy and the environment, where the strategy is contingent on many factors in the firm's environment.

Coherence Theory Positioning

The best fit of coherence is with the RBV, as it starts with a VRIN core capabilities system. It adds the link from competitive advantage to above normal (financial) performance by linking the resources/core capabilities to a market that values these resources/core capabilities. The latter market aspect is a part of the IO theory. In the core capabilities system, DC is incorporated very well. Furthermore, coherence can be seen as a dynamic capability, constantly retaining and improving the fit between the coherence elements. EE and contingency theory are not explicitly incorporated in coherence theory, but their implications can be found in the three coherence elements. It can thus be concluded that the coherence theory is an overarching strategic theory that starts from an RBV view, but extends it by incorporating aspects of the other strategic theories.

2.2.2 Aggregate Findings

The aggregate findings presented here originate mainly from three meta-studies and are discussed to give an overview of the well-established and hardly researched strategic areas. In the last part of this section the Organization-Environment-Strategy-Performance (OESP) model, a meta-theoretical framework incorporating most links in strategy research, is discussed.

Although performed some time ago, the meta-analysis performed in 1986 by Ginsberg and Venkatraman, from a contingency perspective, is still relevant. They note that there are three broad types of contingency variables: environmental, organizational, and performance. These are typically categorized in three levels of organizational strategy: corporate, business and functional. In addition to the broad contingency variable types, four major links of contingency relationships were identified:

(i) The influence of external environment on strategy;
(ii) The influence of organizational variables on the formulation of strategy;
(iii) The influence of performance variables on the formulation of strategy;
(iv) The influence of the chosen strategy on organizational arrangements such as structure, systems and style.

Apart from the individual links shown, some combinations of links are possible as well. Strategy is seen as the central pivot, taking in information from links i, ii and iii to give input to link iv, which in turn influences performance output. Taking the four major links as a starting point many contingency factors have been identified from prior research. They are given in Table 4.

It can be seen that strategy research focused primarily on the relationship between external variables and performance. This does not come as a surprise, since IO is the oldest strategic school and RBV started gaining ground from the 1990's. Within link i, three disciplines are reflected in Table 4: IO (barriers, market structure components and strategic groups), Organisation Theory (environment uncertainty) and marketing (product life cycle, market share position). Most research was performed at business unit level and in multiple industries.

The Ginsberg and Venkatraman meta study still represents the current research focus. Although internal capabilities nowadays receive more attention, most of the research still targets the effects of environmental variables and structural contingency variables (organisational variables) on firm performance.

Table 4 The contingency factors identified by Ginsberg and Venkatraman. Each line represents a single article. Adapted from Ginsberg and Venkatraman, 1985, page 427.

Link	Factor	Strategy level	Scope of analysis	Statistical Significance
i	Market structure	Corporate Strategy	Multiple Industries	No
		Corporate Strategy	Multiple Industries	Yes
	Strategic groups	Business Unit Strategy	Single Industry	Yes
		Business Unit Strategy	Single Industry	Yes
		Business Unit Strategy	Multiple Industries	Yes
		Business Unit Strategy	Multiple Industries	Yes
	Environmental uncertainty	Business Unit Strategy	Multiple Industries	Yes
		Corporate Strategy/Business Unit Strategy	Multiple Industries	Yes
	Product life cycle	Business Unit Strategy	Multiple Industries	Yes
		Business Unit Strategy	Multiple Industries	Yes
	Entry barriers		Multiple Industries	Yes
	Exit barriers	Business Unit Strategy	Multiple Industries	Yes
		Business Unit Strategy	Multiple Industries	Yes
	Market share/growth		Multiple Industries	Yes
	Market share	Business Unit Strategy	Multiple Industries	No
		Business Unit Strategy	Multiple Industries	Yes
	Perceived environment	Business Unit Strategy	Multiple Industries	Yes
	Environmental context	Corporate Strategy/Business Unit Strategy	Multiple Industries	Yes
ii	Perceived need for change	Corporate Strategy	Multiple Industries	Yes
	Managerial characteristics	Corporate Strategy/Business Unit Strategy	Multiple Industries	Yes
iii	Low performance	Business Unit Strategy	Multiple Industries	Yes
		Business Unit Strategy	Multiple Industries	Yes
		Business Unit Strategy	Multiple Industries	Yes
iv	Organization structure	Corporate Strategy	Single Industry	Yes
	Managerial style	Business Unit Strategy	Multiple Industries	Yes
	Internal technology	Corporate Strategy/Business Unit Strategy	Multiple Industries	No

A large meta-analysis was performed by Capon, Farley and Hoenig in 1990. They looked into 320 studies that linked environmental, strategic and organizational factors to financial performance. They classified the relationships investigated in these studies and noted the interaction reported (positive or negative). Statistical analyses were performed with the number of positive and negative correlations reported and conclusions were drawn regarding overall interaction between variables. They provide an excellent graphical summary of their results, which is reproduced in Figure 4.

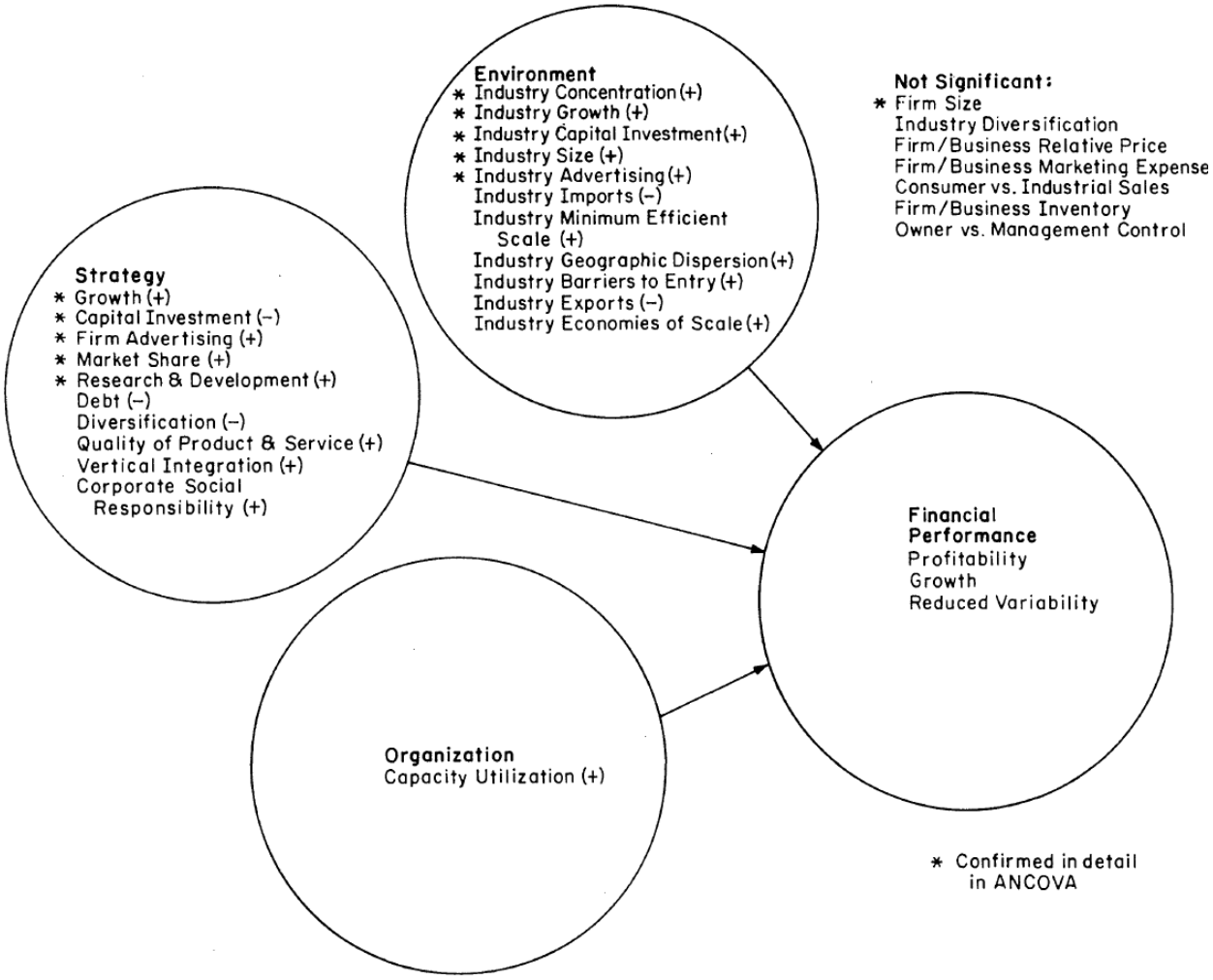

Figure 4 Meta-analysis results of Capon, Farley and Hoenig. Reproduced from Capon, Farley and Hoenig 1990, page 1156.

The first thing to notice in Figure 4 is the lack of organisation-financial performance relationships. The incorporated strategy research has focused primarily on strategic and environmental factors. Although organisational factors nowadays receive more attention, the gap partly remains. Most of the strategic and environmental elements seem logical: growth, market share, research and development, debt, quality, corporate social reponsibility (strategy) and industry concentration, growth, size, scales, entry barriers (environment). Less trivial are the firm and industry advertising that have a positive effect, the diversification discount, the positive influence of vertical integration and geographic dispersion and the negative influence of import and export. Capital investment was found to have a negative influence on an individual firm, but a positive influence on an industry.

Another valuable overview of prior research can be found in the meta-analysis performed by Campbell-Hunt (2000) from an IO perspective. It concludes that 'although cost and differentiation do act as high-level discriminators of competitive strategy designs, the paradigm's descriptions of competitive strategy should be enhanced, and that its theoretical proposition on the performance of designs has yet to be supported', which once more stresses the relevance of this thesis.

The research identified six meta-dimensions of competitive strategy:

- Marketing
- Sales
- Quality reputation
- Product innovation
- Operations
- Market scope

These dimensions were derived from individual elements by cluster analyses. The six-cluster solution provides a clear separation of elements. All elements are associated with one meta-dimension only, except for marketing, which was used in both the marketing and sales dimensions. The elements used originate from literature findings and were previously investigated at least twice with a correlation of 0.5 or higher. They are included in Table 5. Furthermore, the elements were cited by Porter (1980; 1985) and/or Buzzell and Gale (1987) as contributing to competitive strategy.

Table 5 Elements investigated in strategy research at least twice with a correlation of 0.5 or higher. Adapted from Campbell-Hunt, 2000, page 138.

advertising	reputation	product quality	operating efficiency
brand identification	high prices	quality control	unit cost reduction
channel influence	low prices	service quality	modern plant
marketing innovation	new products	procurement	product breadth
promotion	refine products	skilled workforce	customer breadth
sales force	specialty products	manufacturing innovation	

The meta-dimensions contain two to seven statistically significant elements each. Sales and market scope meta-dimensions are narrow; the sales dimension includes advertising, promotion and sales force and the market scope dimension includes product breadth and customer breadth. The operations dimension is much broader, with seven significant elements (Campbell-Hunt, 2000).

Although the study is very informative and summarizes a large amount of research, the usefulness of its conclusions is limited to the differentiation versus cost focus and the notion of the meta-dimensions of competitive strategy.

Conclusion

The meta-studies discussed give a good overview of IO strategy research and a limited overview of RBV research. Throughout the studies discussed many aspects come forward that can intuitively be related to coherence or to the coherence elements. Examples can be found in Table 4, Table 5 and Table 6. In the next section a model that incorporates the different links that came forward in the meta-studies will be explained.

2.2.3 The OESP-model

Farjoun (2002) published about the organic perspective on strategy and the strategic management process. Most relevant for our purposes, he used mechanistic and organic arguments to form the Organization-Environment-Strategy-Performance (OESP) model, a meta-theoretical framework, shown in Figure 5.

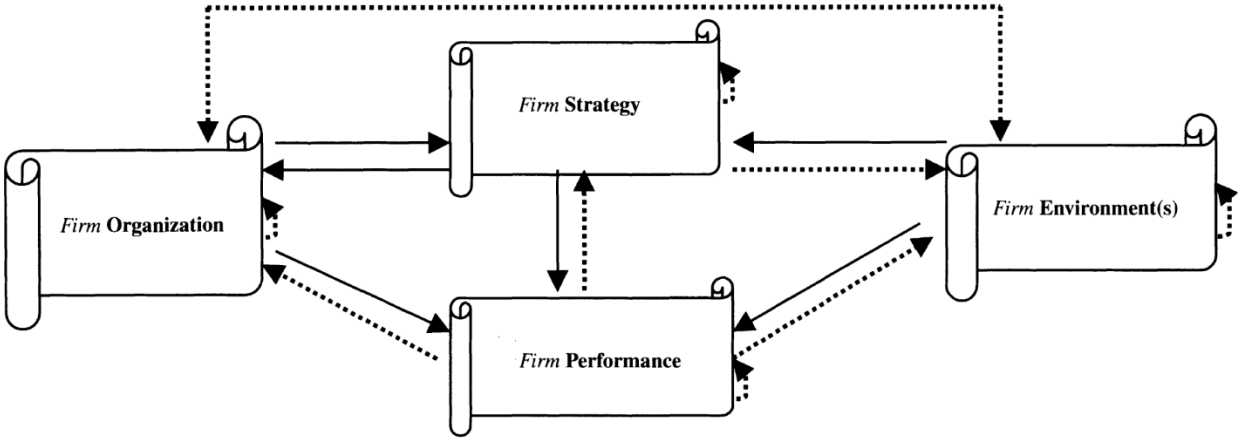

Figure 5 The OESP model. Solid arrows indicate more frequently researched relationships. Dotted lines indicate relationships that received little attention. Adapted from Farjoun, 2002, page 573.

This framework incorporates the commonly accepted groups of factors determining performance. Furthermore it does not exclude any relationships. The relationships that are typically assumed by researchers are marked by solid arrow lines, the less conventional ones are represented by dotted arrows. The latter include for instance the effect of strategy on the environment or the effect of performance on itself. For the latter it makes sense to hypothesize that good (or bad) performance is amplified by the effects it has on the atmosphere in the firm. The effect of performance on strategy makes much sense too.

The OESP-model gives a full overview of the relationships important in strategy research and signals that there are interactions that still need to be addressed properly. Coherence takes into account most of the relationships shown in Figure 5, except for the interactions from performance to other elements.

2.2.4 Strategic Fit

The concept of 'fit' is a central theme to the development of theories in many management disciplines. It has occupied a central role in the formulation of business strategy and business policy, where the concept of "matching" or "aligning" organizational resources with environmental opportunities and threats is strong. Many studies have, either implicitly or explicitly, used the concept of fit. Strategy research studies have explored the roles of a variety of contingency influences on strategy formulation. Most contingency-theory based studies have generally explored the concept of fit by evaluating relationships between two variables, but other views argue for achieving congruence amongst a larger set of elements, which is also referred to as "gestalts" (Venkatraman and Camillus 1984).

Different perspectives of fit

On top of the classification of the general strategy schools, six different perspectives were identified specifically for the concept of strategic fit by Venkatraman and Camillus (1984). It is based on two dimensions: the conceptualisation and the domain of fit. The conceptualisation differentiates between a focus on the content of strategy and the process of strategy formulation. Two categories were defined:

- Strategy as a pattern of decisions made to achieve the optimal match or alignment between the external environment and the organization's internal structure and processes.
- Attempts to specify the strategic actions to be taken as an optimal response to different environmental conditions.

The domain makes the distinction between the environments considered; internal, external or integrated (both internal and external). The two dimensions yield a two by three matrix, which is shown in Figure 6. According to Venkatraman and Camillus (1984) most streams of strategy literature regarding fit can be classified into one of the six cells.

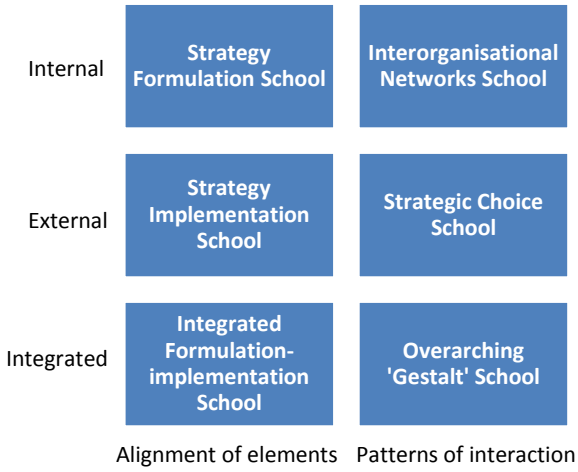

Figure 6 Schools of thought related to strategic fit, sorted by the domain and conceptualisation of fit. Adapted from Venkatraman and Camillus, 1984, page 516.

A short description of each of the schools (Venkatraman and Camillus 1984):

- The Strategy Formulation School believes a firm's performance in the marketplace depends primarily on its (competitive) environment. It focuses on the fit between strategy and external factors.
- The Strategy Implementation School focuses on the alignment between strategy and internal elements, with hardly any direct reference to external influences. An important theme is the strategy-structure fit (structural contingency).
- The Integrated Formulation-implementation School integrates the external and internal focus into a more comprehensive view. It is argued that an organization continually tries to achieve a fit between itself and the environment (external) and between itself and its internal structures and management processes (internal).
- The Interorganisational (Strategy) Networks School recognises that strategies are not purely based on the link between the organization and its environment, but also on the expectations of competitive responses. Therewith it introduces a new level of strategy analysis.
- The Strategic Choice School focuses on the pattern of coordination and interactions amongst internal elements such as structure, size and technology. Here decision making is a strategic choice and not caused by accommodating contingencies. It reflects the important position of the management and the political processes through which such decisions are made.
- The Overarching "Gestalt" School views strategy as an overarching pattern aligning the internal and external elements to the organization. In this context fit is a particular configuration of internal organization and external environment. Organizations are not regarded as autonomous; the chosen strategy has unintended consequences and is affected by the stakeholders that are involved. This supports the argument that the probability of organizational survival increases as the congruence (coherence) amongst environmental, contextual, and structural elements increases.

The best positioning of coherence is probably in the Overarching "Gestalt" School, since coherence incorporates both internal (core capabilities system) and external (product and service portfolio) factors. Furthermore coherence theory tries to achieve a fit between them by applying a specific way to play.

2.2.5 Strategic Alignment

In strategy research strategic alignment can be translated as fitting with each other, or being on the same page. The coherence premium concept can be seen as alignment between the three parts resulting in coherence, which in turn implies a financial premium. In this section an alignment theory of Chorn (1991) is discussed.

Chorn (1991) developed a qualitative framework for creating strategic fit: The ''Alignment'' Theory. It links performance to the alignment amongst four variables:

- competitive situation
- strategy
- culture
- leadership

Each of the variables can take four values, reflecting the following four dominant 'logics':

- production: action and results focused
- administration: control and order focused
- development: creation and change focused
- integration: integration and cohesion focused

The combination of variables and values is given in the matrix in Table 6.

Table 6 Matrix that connects the variables to the dimensions of the alignment theory. Adapted from Chorn, 1991, page 20-23

Dimension\variable	**competitive situation**	**strategy**	**culture**	**leadership**
production	Predictable	operational	rational	growth managers
administration	Repetitive	evolutionary	hierarchical	productivity managers
development	Turbulent	pathfinder	entrepreneurial	creators and builders
integration	Forgiving	protectionist	group	revitalisers

The proposition is that if the same logic is dominant in all the dimensions, this alignment will enhance performance. Although the alignment theory intuitively seems to make sense, unfortunately no empirical research was performed to substantiate the theory. This fact severely hampers its adoption and application. However, the operationalization of alignment is still valuable. It provides a methodology for the assessment of alignment, which can be useful for the operationalization of coherence.

2.2.6 Conclusion

After discussing the aspects of coherence and several related concepts, it is now possible to place the coherence theory within the strategy research field.

To start with, our notion of coherence was not found in the literature directly. The literature is divided in several theoretical schools of thought. The contributions of strategy literature from the different perspectives combined still leave a large (roughly 25-50%) unexplained variance (McGahan and Porter, 1997; Rumelt, 1991). One could ask whether the alignment of the two, coherence, is one of the missing links in strategic theory.

The best fit of coherence is with the RBV, as it starts with a VRIN core capabilities system. It adds the link from competitive advantage to above normal (financial) performance by linking the resources/core capabilities to a market that values these resources/core capabilities. The latter market aspect is a part of the IO theory. In the core capabilities system, DC is incorporated very well. Furthermore, coherence can be seen as a dynamic capability, constantly retaining and improving the fit between the coherence elements. EE and contingency theory are not explicitly incorporated in coherence theory, but their implications can be found in the three coherence elements. It can thus be concluded that coherence theory is an overarching strategic theory that starts from an RBV view, but extends it by incorporating aspects of the other strategic theories.
Looking at the different perspectives of fit identified by Venkatraman and Camillus (1984), the coherence premium concept can best be positioned in the context of the Overarching "Gestalt" School. Coherence incorporates both internal (capabilities system) and external (product and service portfolio) influences. Strategy, a way to play, is a means to achieve the optimal match between the external and internal environment to enhance value creation (Leinwand and Mainardi, 2010).

All in all, many different aspects of the coherence concept can indirectly be found in the literature. Going towards operationalization the Overarching "Gestalt" School combined with Chorn's notion of alignment best represents the coherence premium concept. Although they are useful, the coherence theory framework of Leinwand and Mainardi will be kept, because it is a new concept in which we see potential for measuring strategic fit in a new way.

In the next chapter the implicit links between coherence and literature will be made explicit by mapping previous research onto the coherence premium framework. First, the general coherence framework will be discussed in the upcoming section.

2.3 The Coherence Framework

In this section the model presented in the introduction will be further explained. For this model the following definitions, adapted from Leinwand and Mainardi 2010, 2011a, will be used:

- **Coherence:** Alignment of a firm's core capabilities system, way to play and product and service portfolio
- **Premium:** Above normal financial performance compared to peers (direct competitors)

2.3.1 Extended Model

Coherence can be measured by the three elements it consists of - core capabilities system, way to play and product and service portfolio (Leinwand and Mainardi 2010) - which leads to the initial model as shown in Figure 7.

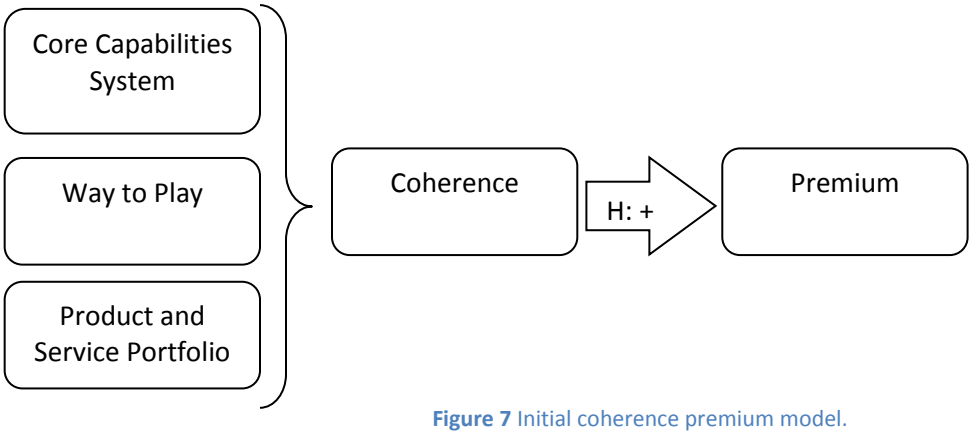

Figure 7 Initial coherence premium model.

2.3.2 Different Perspectives of Fit

The interaction between the three coherence elements can be seen in different perspectives. Venkatraman (1989b) described six perspectives, shown in Figure 8.

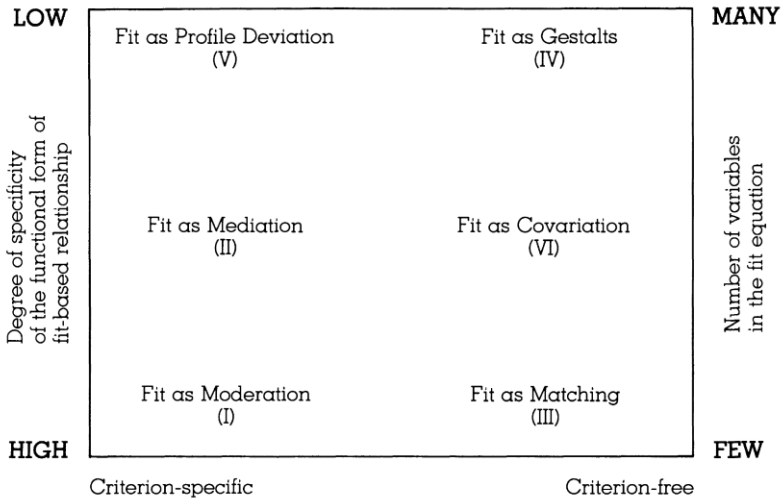

Figure 8 Framework of the six perspectives of fit, ordered by specificity of criterion and form. Reproduced from Venkatraman, 1989b, page 425.

The vertical axis is the degree of specificity or how strict the relationship is tied in a functional form. This axis coincides with the number of variables that is used. The horizontal axis divides the spectrum in either criterion-specific, which tests for a particular relationship, or criterion-free, which does not aim to prove a specific relationship and is thus universally applicable.

Perspective of the extended model

For coherence, only the criterion-specific perspectives are relevant as the aim is to test for a particular relationship, between coherence and premium. Fit as gestalts, covariation and matching are thus inapplicable.

The profile deviation perspective is used frequently to relate combinations between independent variables to the dependent variable. It is appropriate for larger numbers of independent variables. As there is only one independent variable, coherence, profile deviation is discarded.

The model described in Section 2.3.1 translates the scores of the three individual coherence elements into a single coherence score; a combined effect only. This coherence score is related to premium, the

interactions of the individual elements with premium is not taken into account. This is the mediating perspective. If the individual elements are hypothesized to have a direct effect as well as a combined effect, the perspective would be moderating.

In the next chapter literature will be reviewed and mapped onto the coherence premium. This is done to enable the validation, extension and operationalization of the coherence framework in Chapter 4.

3. Focused Literature Review

This chapter explores scientific publications in and around the strategy research field to extend and operationalize the coherence premium framework. The research presented in this literature review is classified and discussed according to the four parts of the coherence premium: way to play, core capabilities system, product and service portfolio and premium. These, and a small section about other relevant factors, can be found in Sections 3.1 to 3.5. Every section discusses relevant literature and maps it onto the coherence elements. Section 3.6 concludes with a graphical summary of the relationships discussed in the previous sections. This procedure is performed to be able to validate, extend and operationalize the coherence framework in the final section of this chapter.

The literature discussed in this literature review is focused on research done in the strategy field, because the coherence premium concept is focused on company strategies. In order to perform the literature research efficiently a four-step methodology was used:

1) Kick-start: relevant and background articles, obtained from experts at Delft University of Technology, as a starting point.
2) Snowball: search through the forward and backward references in the most relevant articles.
3) Ensuring approximate completeness by using the key terms of relevant articles found to search in the most prominent strategy journals (Strategic Management Journal, Academy Management Journal, Academy Management Review, Administrative Science Quarterly and Management Science).
4) General search in Scopus and JSTOR using the key terms of relevant articles.

Although full coverage cannot be guaranteed, the fact that step three added two more articles to the existing set of approximately 120 articles indicates that the majority of relevant articles was found. A subsequent search in Scopus and JSTOR did not yield any new relevant articles.

The findings of articles that were regarded as quantitative and/or very relevant were summarised in detail in the quantitative and relevant literature overview, which can be found in Appendix I.

Throughout the next sections the relevant literature found is mapped onto the coherence premium concept. This approach was taken, because the project started from the coherence premium article. This article is based on decades of experience in strategy consulting work at Booz & Company, indicating that coherence leads to a premium. It gives an interesting overview of the relevance of previously performed strategy research to the coherence premium concept and vice versa. On top of that it will be shown that the majority of previously investigated elements can be mapped onto the coherence framework in a satisfactory manner.

In the end of this chapter Figure 9, which shows the elements in the coherence premium concept of Leinwand and Mainardi, will be filled out. The three coherence elements way to play, core capabilities system and product and service portfolio, are depicted by a Venn-diagram, because this allows a more accurate mapping of relevant literature onto the coherence premium concept. Some research may be on the edge between two elements or may be interpretable in more than one way. The Venn-diagram allows these situations to be mapped accurately.

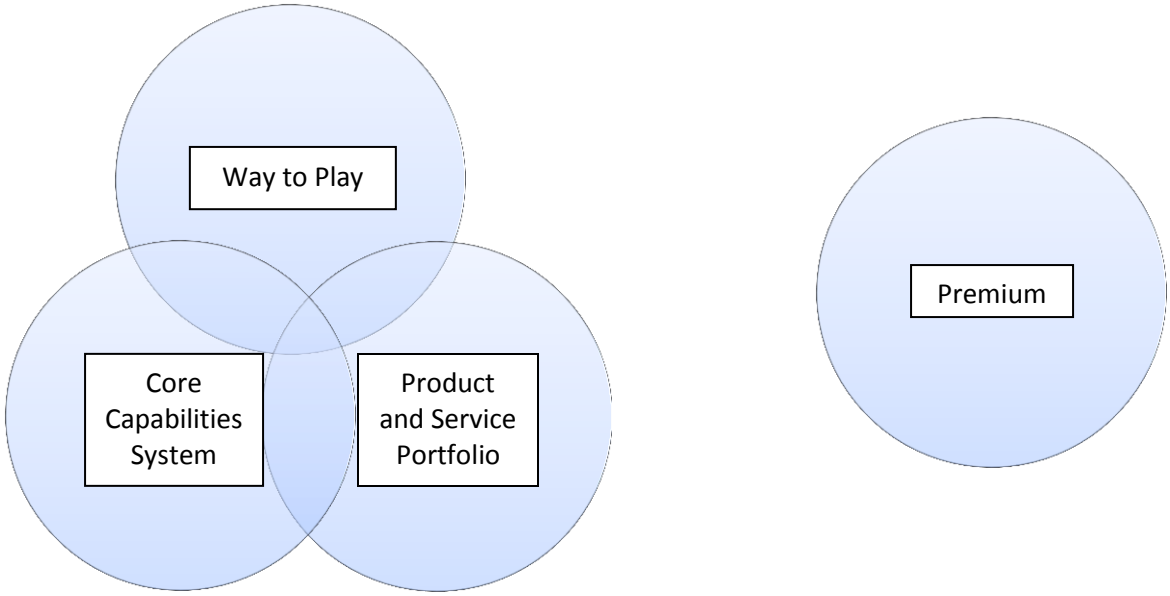

Figure 9 Empty coherence premium framework, which will be filled out in sections 3.1 through 3.5.

3.1 Way to Play

When looking at the way to play, or strategy, firms are often classified in different strategic segments/types by resemblance to the profile of a specific type or group. This section covers segmentations based on costs and differentiation and strategic activity. An alternative approach based on deviation from an ideal profile is discussed as well.

3.1.1 Cost and Differentiation

A basic segmentation is to differ between cost and differentiation focused strategies (Campbell-Hunt, 2000). This differentiation was expanded by Campbell-Hunt to six meta-designs of competitive strategy:

- innovation and operational leadership
- cost economy
- focused quality leadership
- focused quality economy
- sales leadership
- broad quality and sales leadership

This expansion was done by first dividing the studies into cost-focused, differentiation-focused and not-focused studies. Subsequently cluster analyses were performed to give the six meta-designs with the elements from Table 5 in Section 2.2.2. The outcomes of different analyses differed considerably, but the stated classification was regarded as the best option by Campbell-Hunt.

Cost leadership and differentiation were empirically investigated by White (1986) in a study of 69 business units. His results indicate that the fit between business unit strategy and the internal organisation of multi-business companies has an effect on business unit performance. This adds another dimension to the equation: fit with the internal organisation. It could very well be that this fit with the internal organisation is a prerequisite for success with a cost and/or differentiation focused strategy.

More recently, Zott and Amit (2008), who focused on 170 public companies with internet-enabled business models, found indications that novelty-centred business models — coupled with product market strategies that emphasize differentiation, cost leadership, or early market entry — can enhance firm performance. This result confirms that both differentiation and cost leadership are viable strategies to improve performance.

Two well-known classifications related to the separation between differentiation and cost strategies have to be mentioned here. The first and most widely spread is Porter's Generic strategies framework (Porter, 1980), which differs between a differentiation, a cost leadership and a focus (or market segmentation) strategy. Generally spoken, hybrid tactics are not viable according to Porter. However there is one exception: developing innovative products in niche markets, the combination of differentiation with market segmentation, can be very successful.

The other well-known strategy framework was developed by Treacy and Wiersema (1993), who identified three main strategies for value creation:

1) Product leadership: providing the best product and thus related to differentiation.
2) Operational excellence: providing the product at the lowest costs, creating high value for money ratios.
3) Customer intimacy: providing the best total solution, for instance by adding service and personalised products/offerings.

These strategies all lead to added value, but focus has to be on one specific proposition to perform well. While market leaders typically excel at one value discipline, there are companies that have mastered two by resolving the inherent tensions between the two value disciplines.

Both the Porter and Treacy & Wiersema frameworks value the focus on what a company does best; in this aspect it is closely related to the core capabilities system element of the coherence premium concept.

3.1.2 Strategy Typologies

The best-known strategy classification was derived by Miles and Snow (1978). Based on their field studies conducted in the textbook publishing, electronics, food processing, and health care industries, they proposed a strategic typology classifying business units into four groups that exhibit different sets of strategic activities: Prospectors, Analysers, Defenders and Reactors (Miles and Snow 1978).

- **Prospectors** lead industry change, principally by launching new products and identifying new marketplace opportunities.
- **Defenders** find and seek to maintain a secure niche in a stable product area. They stay within a limited product range, focusing more on resource efficiency and process improvements that cut manufacturing costs.
- **Analysers** act as both prospectors and defenders, by both defending positions and sometimes moving quickly to follow opportunities. They are more likely to follow a second-but-better strategy.
- **Reactors** are businesses that lack a consistent strategy, and respond to environmental influences as they arise. They are typically outperformed by the three strategic types with consistent strategies.

Desarbo, Di Benedetto, Song and Sinha (2005) revisited the Miles and Snow typologies by considering three pillars: strategic firm capabilities, environmental uncertainty and performance. This results in a somewhat different classification in four groups. Groups one and two are mixed prospector/analyser groups; group three is a mixed defender/reactor group and group four consists of analysers, defenders and prospectors. Groups differ in their excellence in marketing capabilities, technology capabilities, market linking capabilities, information technologies capabilities, management capabilities. This example illustrates the importance of the Miles and Snow typology 27 years after publication, but at the same time it indicates that successful alternative classifications can be found.

Olson, Slater and Hult (2005) investigated the effects of market turbulence, technological turbulence, formalisation, decentralisation, specialisation, customer orientation, innovation orientation and internal/cost orientation on firm performance (operationalized as the extent to which a firm met expectations, exceeded major competitors, and satisfied top management) for the Miles and Snow strategic types prospector, analyser, low-cost defender and differentiated defender. They concluded that:

- The highest-performing prospector firms have marketing organizations that are characterized by a high number of specialists who operate in a decentralized, informal organization and who place a greater emphasis on customer and innovation orientations.
- The highest-performing analyser firms have marketing organizations whose behaviours are focused on customers and competitors.

The conclusions regarding marketing organisation and orientation drawn here can intuitively be linked to prospectors and analysers based on their definitions. A competition focus will include cost comparisons and defensive behaviour; customer focus is always applicable; decentralisation and innovation are real prospector characteristics.

Slater, Olson and Hult (2006) found that strategic orientation, defined as prospector, analyser, low cost defender or differentiated defender, moderates the strategy formation capability - performance relationship. The strategy formation capability included mission/goal clarity, situation analysis, generation and evaluation of strategic alternatives, and strategy formation process. The sample consisted of 380 firms with over 500 employees from twenty two-digit SIC-code manufacturing/service industries. Apparently strategic orientation has a large influence on the relationship between strategy and performance. As the strategic orientations reflect different type of strategies this makes sense.

Miller and Friesen (1978) identified archetypes of strategy using 31 variables. They identified six successful and four failure archetypes, which are shown in Table 7.

Table 7 Successful and unsuccessful archetypes identified by Miller and Friesen. Adapted from Miller and Friesen 1978, page 921.

Successful archetypes	Failure archetypes
the adaptive firm under moderate challenge	the impulsive firm
the adaptive firm in a very challenging environment	the stagnant bureaucracy
the dominant firm	the headless giant
the giant under fire	the aftermath
the entrepreneurial conglomerate	
the innovator	

Venkatraman (1989a) identified six important dimensions of strategic orientation at the strategic business unit level:

- aggressiveness
- analysis
- defensiveness
- futurity
- proactiveness
- riskiness

These dimensions can be found in many alternative typologies. Taking Miles and Snow as an example, a defender will probably be defensive and not risky, a prospector will be aggressive and proactive and an analyser will be strong in analysis and futurity.

3.1.3 Another Perspective: Profile Deviation

Another frequently used method to look at strategy is by deviation from an ideal profile, which typically is an average of the best performing firms in the industry under investigation. It offers a practical solution to test hypotheses involving many different factors. This is closely related to the way to play, because the specific combination of many different factors comes close to the specificity of the Leinwand and Mainardi (2010) notion of ways to play. The concept of misfits can have a strong managerial impact, as the notion of bad combinations can easily be converted to rules of thumb.

In 1990 Venkatraman and Prescott related environment-strategy co-alignment to performance, using a sample of the PIMS database. Deviations from an ideal profile for each environment[2] were found to have a statistically significant negative relationship with performance[3]. This research shows that profile deviation can be a statistically strong method. Furthermore it indicates that for each environment a single ideal profile (probably with some exceptions) exists.

Naman and Slevin (1993) found that misfit amongst factors[4] diminishes the (survey-based) financial performance. This implies that there are some combinations, the ideal profiles, that are better than others (the misfits). In 2003 Vorhies and Morgan used a sample of 186 US firms in the trucking industry to substantiate that deviation from the ideal marketing organisation profile for a firm's strategic (analyser, prospector, defender) type is negatively correlated with marketing effectiveness (the degree to which the firm achieved its market share growth, sales growth, and market position goals).

All in all the profile deviation method can be valuable for the measurement of specific ways to play, as the method can readily be applied to and come up with ideal profiles from data with many variables. However, this approach is not as direct as a moderating or mediation interaction. When it is applied one can only make a guess about the outcome; instead of validating a theory the results usually lead to new insights on which theory can be built. This is not our approach and therefore the results obtained here are interesting, but the method will not be applied.

3.1.4 Conclusion

This chapter started with the concept of cost versus differentiation strategies. Campbell-Hunt (2000) found six meta-designs of competitive strategy based on these concepts. White (1986) discovered that fit between business unit strategy and the internal organisation has an effect on financial performance. The Treacy and Wiersema framework and Porter's Generic Strategies discussed here are particularly relevant in their call for a focus on what a company does best. There are many strategic typologies, but the Miles and Snow typology (defender, analyser, prospector and reactor) is used most frequently.

[2] The environments were defined as: global exporting, fragmented, stable, fragmented with auxiliary services, emerging, mature, global importing and declining environments.
[3] Performance measured as ROI over four years.
[4] Factors: environmental turbulence, entrepreneurial style, organisation structure, firm's mission strategy.

Profile deviation is another way of measuring fit: the adherence to a specified profile is used to predict performance. This approach can be used in concepts with many variables and approaches the specificity of a way to play.

The individual relationships found in literature are mapped onto the coherence premium concept in Figure 10. The strategy formation capability is a capability and clearly includes strategy as well, therefore it is placed in between way to play and core capabilities system. Marketing is placed in between way to play and product and service portfolio, because products and services are marketed to customers driven by the strategy behind it.Table 8 gives an overview of the same relationships now accompanied by the corresponding authors.

There has been a lot of research on the topic of strategies that a firm can adopt. The way to play coherence element receives strong back-up by this well-established research topic. The specificity of a way to play generally is higher than the strategies in most articles. This probably has to do with the difficulty of measuring strategies at such a specific level and the fact that this reduces data to very small samples for each way to play. Still, the results for more aggregated strategy typologies indicate that a way to play is very important to a company and this justifies its appearance as one of the three coherence elements.

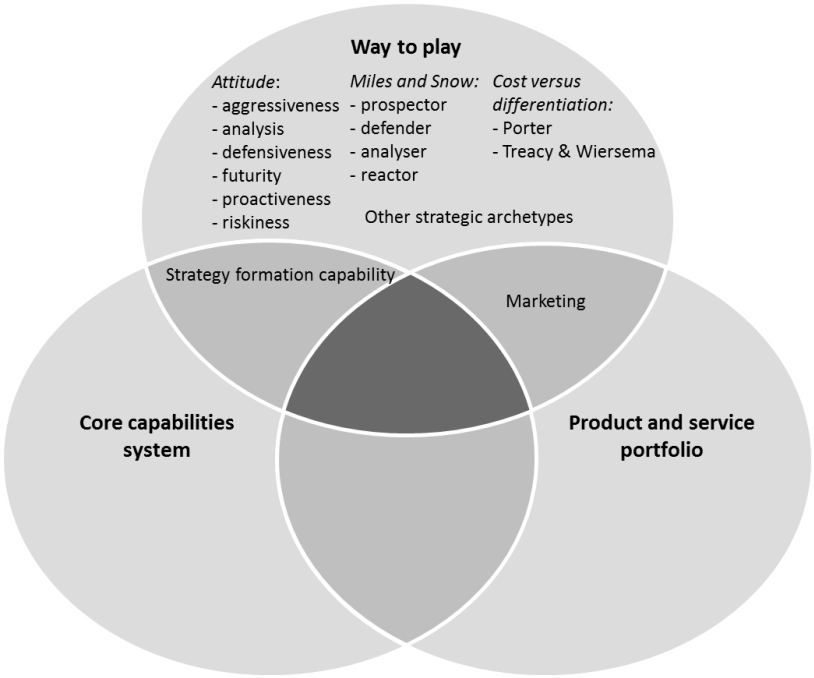

Figure 10 Literature findings way to play mapped onto the coherence premium framework.

Table 8 Overview of way to play literature findings by topic and author(s).

Aspects investigated	Authors
Cost and differentiation	White (1986) Campbell-Hunt (2000) Zott and Amit (2008) Porter (1980) Treacy and Wiersema (1993)
Miles and Snow typology	Miles and Snow (1978) Olson, Slater and Hult (2005) Slater, Olson and Hult (2006)

	Vorhies and Morgan (2003) Venkatraman and Prescott (1990) Desarbo, Di Benedetto, Song and Sinha (2005)
Other strategic archetypes	Miller and Friesen (1978) Desarbo, Di Benedetto, Song and Sinha (2005)
Strategy dimensions (Attitude): Aggressiveness, analysis, defensiveness, futurity, pro-activeness, riskiness	Venkatraman (1989a)
Marketing	Vorhies and Morgan (2003)
Strategy formation capability	Slater, Olson and Hult (2006)
Orientation: customer, innovation, internal/cost	Olson, Slater and Hult (2005)

3.2 Core Capabilities System

Capabilities are typically mentioned explicitly in articles with a Resource-Based View and Dynamic Capabilities perspective. From the latter perspective the focus is on the renewal/adaptation capability, which is important for our notion of capabilities, but it is not necessary to dive into this in much more detail. In this section the first notion of core capabilities, by Prahalad and Hamel in 1990, and a very relevant article by Vickers-Koch are discussed. Subsequently the classification of capabilities will be covered. Furthermore a few relevant empirical findings are presented. In Section 3.2.4 these are mapped onto the coherence premium concept.

3.2.1 The Core Capabilities Concept

The first authors to use the notion of core capabilities (although they call them core competencies) were Prahalad and Hamel (1990). They describe a core competency as a particular strength relative to its competitors that is central to the business of the firm. Core competencies provide the fundamentals for added value and therewith for competitive advantage, which in turn contributes to premium. Core competencies have to fulfil three criteria; they should be non-imitable, re-usable and contributing to end consumers' perceived value (Prahalad and Hamel, 1990). The Prahalad and Hamel core competencies are closely related to Leinwand and Mainardi's (2010, 2011) notion of core capabilities.

In 1995 Long and Vickers-Koch published their view on the use of core capabilities to create competitive advantage. It very much resembles the Leinwand and Mainardi (2010, 2011) interpretation. According to Long and Vickers-Koch capability-based organizations find the best fit between a firm's resources, business conditions and markets by defining their resources in terms of the capabilities they have developed for adding value for their customers and other stakeholders. Threshold capabilities are necessary to compete. Leinwand and Mainardi call these 'table stake capabilities'. The capabilities a business relies on for its competitive advantage are core capabilities. Long and Vickers-Koch (1995) make an additional distinction between:

- Critical core capabilities, competencies and processes that provide the current competitive advantage.
- Cutting edge core capabilities, competencies and processes that will provide future competitive advantage.

This competitive advantage results from the company's ability to balance its capabilities with its opportunities. This can be achieved at four levels:

1. Existing core capabilities and market segments are examined and opportunities to use those capabilities to create new products and services in those market segments are identified.
2. Existing market segments can be analysed to determine what new core capabilities could be developed to protect the firm's position in those markets.

3. New products and services are created by redeploying or recombining existing core capabilities to satisfy new market segments.
4. At the highest level of innovation, a company can search for the most promising new market segments and try to develop new core capabilities to compete there.

At any level, core capabilities are the most critical and most distinctive resources a company possesses. They are difficult for competitors to copy because they are based on knowledge, skills, and processes developed over time into workable combinations within the context of a particular organizational setting. (Long and Vickers-Koch, 1995).

The similarities between the definitions of Long and Vickers-Koch and Leinwand and Mainardi are very large. But there are subtle differences: whereas Leinwand and Mainardi stress the importance of an integrated system of mutually reinforcing core capabilities, Long and Vickers-Koch focus on balancing capabilities and opportunities in a four level construct.

3.2.2 Capability Classifications

Capabilities and competencies were not frequently linked to performance directly, but a lot has been written about their importance and connections. As the number of (core) capabilities is potentially unlimited, a systematic classification is important in dealing with capabilities. This section first discusses how resources can become capabilities, which in turn can lead to sustained competitive advantage. Subsequently the classification of capabilities will be discussed, using three methods: the American Productivity and Quality Center (APQC) Process Classification Framework, the functional classification and the value chain analysis.

Relevant Definitions

The definitions of resources, competences, capabilities and competitive advantage are often different, depending on the source. Below, the definitions adopted by Grant (1991, page) are cited. The Leinwand and Mainardi notion of core capabilities comes closest to the combination of the distinctive competencies and strategic assets definitions used by Grant. The competitive advantage definition of Grant is close to a successful way to play.

- *Resources:* assets of a firm which can be divided into three broad categories, i.e., tangible, intangible and human resources;
- *Capabilities:* firm's skills at coordinating its resources and putting them to productive use;
- *Distinctive competencies*: firm specific strengths that allow a firm to differentiate its products from those offered by rivals and/ or achieve substantially lower costs than rivals by achieving superior efficiency, superior quality, superior innovation and superior customer responsiveness;
- *Strategic Assets:* Difficult to trade and imitate, scarce, appropriable and specialized resources and capabilities;
- *Core competencies:* Pool of experience, knowledge, systems, etc. that exists elsewhere in the firm, deployed to reduce the cost of creating strategic assets and to reduce the time taken to create strategic assets;
- *Competitive advantage:* implementing a value creating strategy not simultaneously being implemented by any current or potential competitors;
- *Sustained competitive advantage:* when these other firms are unable to duplicate the benefits of this strategy.

According to Grant (1991) capabilities are one side of the medal; the other side consists of industrial factors. Combining them into an appropriate strategy leads to competitive advantage.

Capabilities Classification

There are more ways to look at the different capabilities that exist. Leinwand and Mainardi adopted a very detailed view of a core capability. There are thousands of these capabilities. An alternative is to look at categories of capabilities. This results in a more limited set of capabilities, which lacks the specificity of the Leinwand and Mainardi approach. However, this can be advantageous to give an overview or classification of capabilities and to measure capabilities at a more aggregated level. Capabilities can for instance be labelled by functional department, strategic department, steps from in-to output etc.

In this section alternatives and aggregations will be discussed. The APQC process classification framework is treated first, followed by the value chain and functional classifications.

APQC Process Classification Framework

American Productivity and Quality Center (APQC) is an organization specialised in benchmarking studies and research. For industry and cross-industry comparisons, APQC leverages a process classification originating from IBM.

The APQC process classification framework has four levels of capabilities. Level one is an organizing structure and the items here are not capabilities themselves. Level one differentiates between operating processes and management and support processes. Generally, the L1 models follow function, i.e. research and development, marketing, sales, etc. An example of a level one is given in Table 9.

Table 9 Level one of the APQC process classification framework. Adapted from Architecting Value.

Operating Processes	Management and Support Processes
1 Develop Vision and Strategy	6 Develop and Manage Human Capital
2 Develop and Manage Products and Services	7 Manage Information Technology
3 Market and Sell Products and Services	8 Manage Financial Resources
4 Deliver Products and Services	9 Acquire, Construct and Manage Property
5 Manage Customer Service	10 Manage Environmental Health and Safety
	11 Manage External Relationships
	12 Manage Knowledge, Improvement and Change

Level two capabilities provide much greater precision, but are still organising structures. In level three and four it becomes very specific and the items there can be regarded as Leinwand and Mainardi (core) capabilities. Level three is a capability – '*a business outcome unconstrained by how it is implemented by people, process, information, or technology*'. Level four in the APQC model is generally more process-oriented. Usually there are several hundreds of items for an enterprise at L3. This provides considerable precision without losing all the overview (Architecting Value).

Value Chain Analysis

An alternative is to take the value chain to classify capabilities. As an example Porter's value chain is shown in Figure 11. A distinction is made between primary activities and support activities.

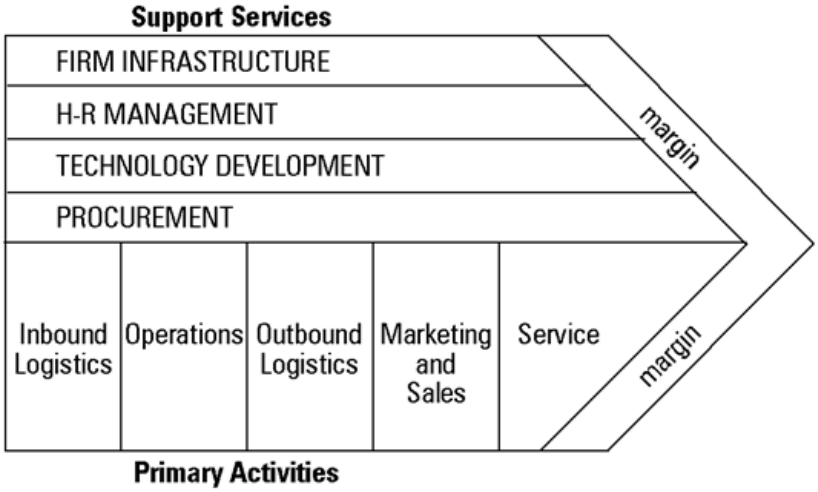

Figure 11 Porter's Value Chain, which differentiated between primary activities and support services. Both can contribute to the margin made on the products or services sold. Adapted from 12manage.

The primary activities follow the offered products/services in all the steps from entrance to exit. Supporting activities are supportive of all the primary activities, like information technology, human resource management etc.

Functional Classification

Grant (1991) uses the functional classification as well. In this approach the capabilities can be found in the functional departments and it is easy to use, as the functional classification is already present and used frequently in most organisations. An example of a functional classification is given in Table 10.

Table 10 An example of the functional classification method. Adapted from Grant 1991, page 136.

Functional Area	Capabilities
Corporate Functions	Financial control
	Strategic management of multiple businesses
	Strategic innovation
	Multidivisional coordination
	Acquisition management
	International Management
Management Information	Comprehensive, integrated system linked to managerial decision-making
Research and Development	Research
	Innovative new product development
	Efficiency in volume manufacturing
	Continuous improvement in operations
	Flexibility and speed of response
Product Design	Design capability
Marketing	Brand management
	Promoting reputation for quality
	Responsiveness to market trends
Sales and Distribution	Effective sales promotion and execution
	Efficiency and speed of order processing
	Speed of distribution
	Quality of customer service

Conclusion

The most appropriate classification depends on the level of analysis. The APQC Process Classification Framework and functional classification are best suited for a very deep and more aggregated level respectively. Additionally, the value chain analysis can be used to look for and identify capabilities from another perspective. Integrating the functional and value chain approaches, a set of overarching core capability categories was developed, as shown in Table 11.

Table 11 Overarching categories of core capabilities based on the functional and value chain classifications.

- **Supply chain management**
  - Distribution
  - Logistics
  - Sourcing
  - Channel activation
  - Supplier/customer relationships
- **Marketing**
  - Brand management
  - Advertising
- **Portfolio management**
  - Acquisitions
  - Divestitures
- **Manufacturing**
  - Production
  - Engineering, basic and applied
  - Research and development
- **Selling**
  - Sales force
  - Pricing
  - Go to market strategy
- **Financial management**
  - Working capital
  - Performance measures

3.2.3 Empirical Research Examples

In this section several relevant empirical articles, typically performed in specific areas, that provide one or more capabilities of a firm is presented. These give an indication of research performed on this topic.

Intangible resources

Hall (1993) researched intangible resources and capabilities. He argues that sustainable competitive advantage results from relevant differentiating capabilities. He also argues that regulatory and positional capabilities are related to intangible assets like intellectual property, contracts, data, networks, know-how, reputation and culture. He proposed a framework linking intangible resources to capabilities and tested it with six executives of successful companies. Intangible assets can definitely be part of core capabilities, but this research cannot reliably indicate its importance due the small panel of six executives.

Strategy-making Process Capability

Using data on five dimensions of perceived performance from 285 top managers, Hart and Banbury (1994) found that firms with high process capability, using multiple strategy-making process modes simultaneously, outperform single-mode or less process-capable firms. This can be seen in the light of core capabilities. The number and level of strategy-making process modes used is a core capability.

New Product Development

Investigating New Product Development (NPD) Acur, Kandemir and Boer (2012) coupled strategic planning, innovativeness, environmental munificence and technological change to technological, market and NPD marketing alignment and eventually to NPD performance. Strategic planning and innovativeness were found to have a statistically sound and positive relationship with all three types of alignment. NPD marketing alignment and, less pronounced, technological alignment were found to be positively related to NPD performance. NPD can be very important for firms with a differentiation

strategy. Strategic planning and innovativeness are indicated to improve NPD and can thus be regarded as underlying capabilities of NPD alignment and performance.

3.2.4 Conclusion

The concept of core capabilities was founded by Prahalad and Hamel in 1990. They state that a core capability has to fulfil three criteria: it should be non-imitable, re-usable and contributing to end consumers' perceived value. Long and Vickers-Koch adopt a view on core capabilities very consistent with Leinwand and Mainardi. Apart from differentiation between threshold and core capabilities, they distinguish critical core capabilities, competencies and processes that provide the current competitive advantage, and cutting edge core capabilities, competencies and processes that will provide future competitive advantage. Further additions by Long and Vickers-Koch include a focus on balancing capabilities and opportunities in a four level construct. The importance of an integrated system of mutually reinforcing core capabilities, as stressed by Leinwand and Mainardi, is not mentioned explicitly. Little empirical research had been performed on the effect of core capabilities on firm performance. The notion of core capabilities is only a specific part of the set of internal factors in strategy research. Almost any resource can be(come) (part of) a core capability, but a core capability cannot be defined as a specific resource. Few scholars have looked at resources with the specific capabilities lens.

The number of (core) capabilities is potentially unlimited. Therefore a systematic classification is important in dealing with capabilities. Three methods were identified; the APQC process classification framework can be used for extensive and deep investigations, the functional classification is the simplest to use and the value chain analysis adds a different point of view. These methods were combined into an overarching classification of capabilities, as shown in Table 11.
Table 12 gives an overview of the same relationships, accompanied by the corresponding authors.

Several examples of investigated relationships between specific capabilities and premium are mapped onto the coherence premium concept in Figure 12. Strategic planning is a capability, but strategy-making is the interaction between strategic choices and its capability dimension. Therefore it is added in between way to play and core capabilities system, together with innovativeness, which is a capability and a strategic choice. New product development can be a very strong capability and the product part clearly links it to the product and service portfolio, which explains its positioning in between them.

Relevant articles concerning capabilities were found, but their number is limited. Their appearance starts about two decades ago, which partly explains this. Additionally, there might be sources that use other terms for capabilities, which made them slip through our radars. Even if this is the case it must be concluded that the research field is still young and small. There is a lot of unleveraged potential in combining insights of different capabilities practitioners, like Booz & Company. However, their data is often confidential, limiting the availability. The lack of findings on core capabilities could also be due to the specificity of the concept, which can make the measurement problematic. This would strengthen the Leinwand and Mainardi (2010) notion of core capabilities. All in all, the findings are compatible with and substantiating the core capabilities coherence element. Therefore this element is backed up by literature, which explains and confirms its importance as one of the three coherence elements.

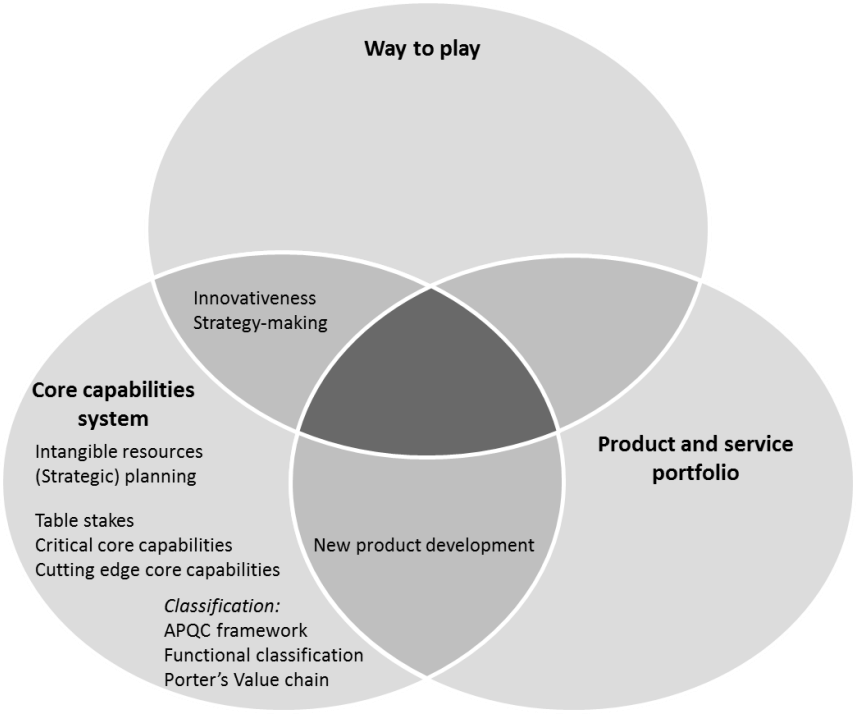

Figure 12 Literature findings core capabilities system mapped onto the coherence premium framework.

Table 12 Overview of core capabilities system literature findings by topic and author(s).

Aspects investigated	Authors
Concept of core capabilities	Prahalad and Hamel (1990) Long and Vickers-Koch (1995)
Intangible resources	Hall (1993)
Strategy-making process	Hart and Banbury (1994)
New product development Innovativeness Strategic Planning	Acur, Kandemir and Boer (2012)

3.3 Product and Service Portfolio

The product and service portfolio is the interface between a company and its environment; through its products and services it is influenced by market conditions. It is therefore closely related to the research of external influences. However, for coherence it does not matter in which industry or industries a company competes. Therefore the effect of industry on coherence is covered with the other relevant factors in Section 3.4. For coherence it is very important that the system of core capabilities is applied to all the segments in which it is active. This becomes harder when a firm competes in more segments. That is why this section gives a discussion of diversification, an important concept related to the width of the product and service portfolio. Finally, the relationships identified from the literature will be mapped onto the coherence premium concept.

3.3.1 Market Influence

For coherence in its essence it does not matter in which market or in how many markets a company is active. The relative coherence compared to competitors is most important. However, ways to play are specific and their success depends on market characteristics. Therefore the specific way to play is dependent on the market, but the successful combination of core capabilities, way to play, and product and service portfolio is what affects premium.

It is important to note the differences between two meanings of a 'market':

- The industry or the segment, which by its specific activities and most common business models partly determines the normal level of premium measures in the market. This is an effect, which will have to be taken into account, because it is an alternative cause for differences in premium. This is covered in Section 3.4.
- The market in the sense of the competitors, customers, rules and regulations, product offerings, stakeholders, relationships, important patents etc. This is the market definition Leinwand and Mainardi (2010) use. Given the characteristics of a market it is possible to combine the strengths, in the form of core capabilities, with the demands of customers (product and service portfolio) by applying a successful way to play.

Miller and Friesen (1983) investigated the environment – strategy-making link, hypothesizing that increases in environmental dynamism, hostility and heterogeneity should relate to specific changes in the analysis and innovation aspects characterising the strategy-making activity amongst two distinct firm samples. Most of these relations were found to be much stronger in successful than in unsuccessful firm samples. This indicates that successful firms are in closer touch with their environment and act accordingly.
McArthur and Nystrom (1991) used 109 firms in 35 manufacturing industries to look into the moderating effect of environmental conditions on the strategy-performance relationship. Their analyses show that environmental dynamism, complexity, and munificence each moderates the form of strategy-performance relationships. They found that inventory turnover, slack and plant newness were positively related to performance[5], whereas capital intensity was negatively correlated with performance.

3.3.2 Diversification

Diversification is a measure for the width of the product and service portfolio. Diversification is an important theme in research and there has been a lot of debate about whether or not diversification destroys value. Generally a small diversification discount is often found, but as there are enough individual exceptions and contradictory results, the final conclusion cannot be drawn yet. (Wernerfelt and Montgomery 1988; Servaes, 1996; Campa and Kedia 2002; Villalonga 2004; Choe and Yin, 2009) Recently Impink and Reppenhagen (2012) conclude that reported segments do not reliably measure diversification and that this is likely to introduce a bias towards a (larger) diversification discount. They propose the number of subsidiaries as an alternative proxy for diversification. This exemplifies the state of the current debate.

Markides and Willliamson (1994) argue that related diversification can be a competitive advantage if it assists the firm in accumulating non-tradable, non-substitutable assets efficiently. Therefore related strategic assets prevail over related markets. By transferring core competences between its strategic business units, a corporation can accelerate the rate and lower the cost at which it accumulates new strategic assets. The related diversification could have been labelled capabilities-driven diversification by Leinwand and Mainardi. By using the same core capabilities in all product categories a company can add a lot of value, whilst being able to divide the costs of the core capabilities over the entire product range.

Diversification is often measured by the number of industries or segments in which a company competes. An often used standard industry classification is the SIC-code (Standard Industrial Classification code) classification, which is a US government system for classifying industries by a four-digit code. The SIC codes are grouped into industry group, major group and division. The first two digits

[5] Measured by the Return On Investment (ROI).

indicate the major group, the first three digits indicate the industry group. SIC-codes were frequently used as a measure for diversification in the strategy literature discussed in this thesis.

3.3.3 Conclusion

The product and service portfolio is the link to the outside world, to the markets and industries. It is the market where the link between core capabilities and customer value is made. It is important to differentiate between the normal values of financial metrics in an industry, for which results should be compensated, and the specific market characteristics, which define value-adding capabilities and successful ways to play. Product and service portfolio in coherence theory is about the latter. The investigated relationships discussed are mapped onto the coherence premium concept in Figure 13.

Table 13 gives an overview of the same relationships now accompanied by the corresponding authors.

Miller and Friesen (1983) found indications that successful firms are in closer touch with their environment and act accordingly. McArthur and Nystrom (1991) looked into the moderating effect of environmental conditions on the strategy-performance relationship. Their analyses show that environmental dynamism, complexity, and munificence each moderates the form of strategy-performance relationships.

Another representation of service and product portfolio is diversification. Although still topic of debate amongst scholars, generally speaking literature findings indicate that diversification comes with a small discount. However there are many exceptions to this rule. Diversification can be very beneficial, if the existing core capabilities are all used for the new product category and all add significant value to it.

Product and service portfolio is probably the most difficult of the three coherence elements, because it is a specific part of the 'market'. Therewith it differentiates from the IO and contingency theory notion, as product and service portfolio is about the internal market characteristics and the fit with the other coherence elements and not about the effects of being in a different industry, having a different firm size or being diversified. Indeed the coherence theory does not include diversification under the product and service portfolio element. Although the arguments explaining that this essentially should make no difference are understandable, the difficulty of achieving coherence with rising diversification is recognized from practical considerations. Therefore diversification is added here. Apart from diversification there is little evidence in the literature that uses similar definitions. There is, however, a lot of work done on for instance the effects of competitive forces. All in all, the product and service portfolio dimension is quite different from the views expressed in the literature findings. This makes it the most vulnerable of the three coherence elements, because it cannot directly be substantiated by previous research findings.

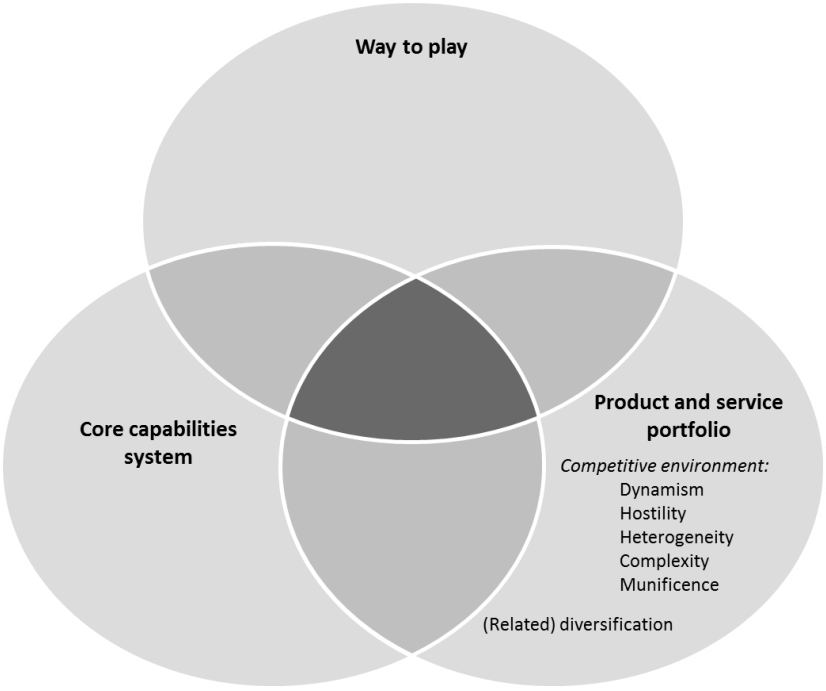

Figure 13 Literature findings product and service portfolio mapped onto the coherence premium framework.

Table 13 Overview of product and service portfolio literature findings by topic and author(s).

Aspects investigated	Authors
Competitive environment: Dynamism, hostility, heterogeneity, complexity and/or munificence	Miller and Friesen (1983) McArthur and Nystrom (1991)
Diversification	Wernerfelt and Montgomery (1988) Servaes (1996) Campa and Kedia (2002) Villalonga, (2004) Choe and Yin (2009) Markides and Willliamson (1994)

3.4 Other Factors Affecting Premium

Apart from the elements that have thus far been mapped onto the coherence premium framework, there are several variables that do not explicitly fit within the concept. These were therefore not included in the Venn-diagram. These variables include environmental influences and structural contingency factors. The former are dependent on the industry a firm is active in and the latter are properties that do not change easily or rapidly or cannot be influenced by a firm in the short or medium term. In this section these other relevant factors will be discussed and in the last section their relevance to this research is explained. The method to incorporate them into the research is explained here as well.

3.4.1 Structural Contingency Variables

Examples of structural contingency variables that can be found in the research from previous paragraphs include size, age and structure[6]. Especially age is a property that cannot be influenced. Although they can be important firm properties, they will change on the long-term due to present,

[6] For instance formal/informal, centralised/decentralised or locus/span of control.

historic and future decisions and strategy making. When creating strategy, structure will be taken into account. Structure is correlated with size, because size enables or disables the use of certain (centralised) structures. An example of research on structural contingency variables about organisational alignment is included below.

Powell (1992) researched organisational alignment as a competitive advantage in two consumer goods manufacturing industries, looking at internal-structural fit, size-structure fit, size-planning competency fit, industry-structure fit, industry- planning competency fit and locus of control. These were coupled to supernormal profits (3-year profitability) as a performance measure. Internal-structural fit, size-structure fit and size-planning competency fit were found to be positively correlated with performance and statistically significant. Industry-structure fit, industry-planning competency fit and locus of control were found to be negatively correlated to performance, where only the locus of control achieved statistical significance.

The planning competency from this research could be a core capability. The conclusion that a fit with the size of the company gives better results is logical, since large companies demand different planning capabilities as compared to small companies. The rest of the variables are not easily related to the coherence premium concept. Size, structure, internal and locus of control are all structural contingency factors. These factors were extensively researched in the search for the optimal organisation for a company.

3.4.2 Environmental Influences

In this section the other meaning of market is briefly discussed. It is about the effect of differences between industries on the normal level of performance metrics for these industries.
In 1986 Prescott demonstrated that environments, measured by general characteristics of market structures, moderate the strength but not the form of strategy variables - performance relationships. The market thus has a large influence on performance. This was confirmed by Wernerfelt and Montgomery in 1988, when they found that industry effects account for most of the explained performance variance, using Tobin's q (market value / book value) as a measure. With this result one could conclude that the key to success is to choose the industry to compete in wisely. However, the best industry to be in will change over time, as attractive industries will attract more newcomers and therewith more competition.

One year later Hansen and Wernerfelt investigated determinants of firm performance using the Compustat database. They found that industry profitability, relative market share, emphasis on human resources and emphasis on goal accomplishment were positively correlated with 5-year averaged return on assets. Size (natural logarithm of total assets) was negatively correlated with this return.
More recently Ruefli and Wiggins (2003) investigated the degree to which firm performance is associated with industry or corporate factors using the Compustat database and five-year ROA as a performance measure. Results indicated that corporate factors were an order of magnitude better predictors of business unit profit position and industry factors were found to be non-significant predictors. This result is sharply contrasted to the findings presented earlier in this section.

Some well-known empirical work with large datasets was done by Rumelt (1991) and McGahan and Porter (1997) regarding the effect of industry, business unit and corporate parent on performance. They react on each other's work and disagree, which strikes me as being quite strange as both articles state that the largest variance is explained by the business unit, on a distance followed by the industry.

What can be concluded is that a lot of research has been done on external influences. Industry and business unit clearly have a large influence on performance and other elements like market structure, market share and corporate parent have some influence as well. However, caution has to be taken since not all results point in the same direction, which could indicate that the research method chosen has implications for the outcome.

3.4.3 Relevance

The other factors influencing premium are highly relevant, because they provide alternative explanations for premium. This means that if they are not taken into account the relationships found between coherence and premium might not be caused by coherence at all. Therefore the results should be corrected, or at least checked, for the effect of these factors.

With our data it is possible to take into account size, industry (by SIC-code and by French classification) and fiscal year. After performing exploratory analyses, the effect of these factors on the relationship found will be assessed. If the remaining dataset is large enough, the correction could be performed for more or all of these factors simultaneously.

3.5 Premium

Premium is the end measure in the coherence theory. It is defined as *above normal financial performance compared to peers*[7].

First, the question why coherence leads to a premium should be answered. This is done in the first section. In Section 3.5.2 some of the many different premium measures that can be found in literature are presented. Research assessing the validity and correlation between various financial metrics is discussed here as well. In the end the premium balloon is filled up with the possibilities.

3.5.1 Relationship between Coherence and Premium

In this section the question why coherence leads to a premium will be answered by looking back at the literature research performed thus far. The explanation why coherence leads to a premium was not found in literature.

The Coherence Premium

Leinwand and Mainardi (2010) give four reasons why coherence leads to a premium[8]:

- Coherence strengthens a company's competitive advantage by focus on their capabilities.
- Coherence focuses strategic investment.
- Coherence produces efficiencies of scale.
- Coherence creates alignment between strategic intent and day-to-day decision making.

This is strengthened by Booz & Company's consulting experience, which shows the coherence premium at work with their clients.

Other Literature Findings

The coherence elements are content-wise not new to the strategy research field. The coherence theory combines aspects of often researched strategy factors influencing (financial) performance, including resources, capabilities, organisation, strategy and environmental factors.

Research aimed at fit between strategy, environment and organisation can be put in the perspective of the coherence theory. Product and Service Portfolio represents environment in the sense of customer

[7] Peers: direct competitors,
[8] More elaborately discussed in Section 2.1.4

needs, competitiveness and other market characteristics; the core capabilities system represents the organisation in the sense of its strengths and weaknesses, resources and capabilities; way to play is the specific strategy or business model that combines the environment and the organisation. Capon, Farley and Hoenig (1990) showed that results found for many aspects within strategy, organisation and environment have been found to have a (in most cases) positive influence on financial performance (profitability, growth and reduced variability). In his general framework, Farjoun (2002) includes strategy, organisation and environment as meta-factors influencing performance as well.

Another type of research that can be related to coherence is research using profile deviation. It tests whether specific combinations amongst many strategy, organisation and environmental factors are more successful than others. Venkatraman and Prescott (1990) found that profile deviation significantly diminishes return on investment. Naman and Slevin (1993) show that misfit between mission strategy, structure, entrepreneurial style and environment decreases financial performance. Desarbo, Di Benedetto, Song and Sinha (2005) revisited the Miles and Snow typologies to find interrelations between strategic types, capabilities, environment and performance.

Additionally, there are findings that indicate the relevance and relationship of the coherence elements individually with premium. Some of them will be presented briefly here.

Core Capabilities System
From the Resource Based view, many researchers hypothesize a positive interaction between resources and premium or performance. More specifically Pramel and Hamalad (1990) and Long and Vickers-Koch see core capabilities as the main drivers of performance. Chmielewsky (2007) found that resource orientation leads to a premium (higher return on capital and return on sales) for consumer goods companies.

Way to Play
Miles and Snow (1978), Miller and Friesen (1978), Porter (1980), Treacy and Wiersema (1993) all identify several general ways to play. The successfulness of these ways to play is the topic of many papers. Olsen, Slater and Hult (2005) found that the best performing prospector and analyser firms have different marketing organisations. This implies that there are more successful ways to play in a market. Zott and Amit (2008) showed that several successful couplings were possible between novelty-centred business models and product market strategies for public companies with internet-enabled business models.

Product and Service Portfolio
Markides and Williamson (1994) found that related diversification can be a competitive advantage. Furthermore they state that by transferring core competences between its SBUs, a corporation can accelerate the rate and lower the cost at which it accumulates new strategic assets. By using the same core capabilities in all product categories a company can add a lot of value, whilst being able to divide the costs of the core capabilities over the entire product range.

All in all many (indirect) cues can be found that indicate the positive effect of coherence (elements) on premium measures. Indications for a negative effect were not found. Therefore, the hypothesis is that coherence is positively related to premium. However, it remains difficult to find reasons, in addition to Leinwand and Mainardi (2010), which explain why coherence leads to a premium. It is concluded that this is at least partly caused by the state of the research field itself. The explanation for the largest part

of firm success has not yet been found and most research is aimed at improving this. Therefore little attention is paid to the reasoning behind the factors of (limited) influence.

3.5.2 Premium Measures

The different measures found are summarised in Table 14. To improve the clarity of presentation a distinction is made between capital efficiency, margin, growth and other measures. Most publications reviewed in this literature review use self-reported performance, usually determined by several questions in a survey. Within the capital efficiency measures, return on assets is used most often. EBIT margin (or a derivative thereof) is often used as a profitability measure, but profitability is mostly self-reported. Concerning growth the usual parameters are sales growth and market share (growth). Other measures include self-reported performance and reduced variability as well as numerical measures like total value added and market to book value.

Leinwand and Mainardi (2010) define premium as the EBIT margin, which is equal to EBIT divided by net revenue, over a five-year period. The question remains whether this is the best measure for premium. Another question is how much time it takes for (in)coherence to be visible in financial results. It could also be beneficial to use several premium measures to triangulate the conclusions and to reduce the dependence on bias in a single metric.

Table 14 Premium measures used in literature, sorted in four categories; capital efficiency, margin, growth and other.

Capital Efficiency	Margin	Growth	Other
Return on Investment	EBIT margin	Sales growth	Performance (qualitative, survey)
Return on Assets	Profitability	Growth/share	Market value / book value
Return on Capital		Market Share	Reduced variability

Measure Characteristics

The accuracy and robustness of performance measures are important to consider, before making a choice for one or more premium measures. Several scholars have made comparisons between measures and tested the validity of the measures. Three investigations are interesting to mention.

Chakravarthy (1986) concluded that no single profitability measure seems capable to differentiate between seven excellent and seven non-excellent companies, rated by Fortune, in the computer industry. He argues that financial metrics are just a part of performance and suggests adding stakeholder satisfaction to the equation. Definitely there is more to performance than pure financial elements, but in the end the company exists to generate cash flows. Therefore, financial metrics are still very valid as an end measure if they represent the generation of cash flows. The stakeholder satisfaction will have to translate into monetary value. Additionally, from a practical perspective, adding stakeholder satisfaction is not a simple task, because there can be many and they can be very different from each other.

Naman and Slevin (1993) tested performance primarily with a survey instrument, but checked it with return on sales (net income after taxes divided by gross sales) from sales data and income figures shared by half of the respondents. They reported a small correlation of 0.295 between survey based performance and return on sales, with significance at the 0.05 level. It shows that the (relative) bias from surveys is limited. However, the results are not strong enough to change the view on surveys regarding their subjectivity to biases.

Results from Varaiya, Kerin and Weeks (1987) indicate that profitability and growth have a positive influence on shareholder value. Profitability was found to be stronger correlated with premium than growth.

It is valuable to use more than one premium measure to compensate for individual biases, but the premium measures should be inherently different from each other, preferably measuring a different aspect of premium.

3.5.3 Summary

First the rationale for the positive relationship between coherence and premium was expanded. It can be concluded that most of the literature research discussed in this chapter provides at least a partial argument for the positive relationship.
Then several possibilities for premium measures were discussed. These are summarised in Figure 14. Reduced variability and performance have been excluded, because they are qualitative and hard to measure. It was found that several measures are related to each other, the classification in the figure was done on this basis. In the methodology chapter the end measure(s) will be chosen based on this list.

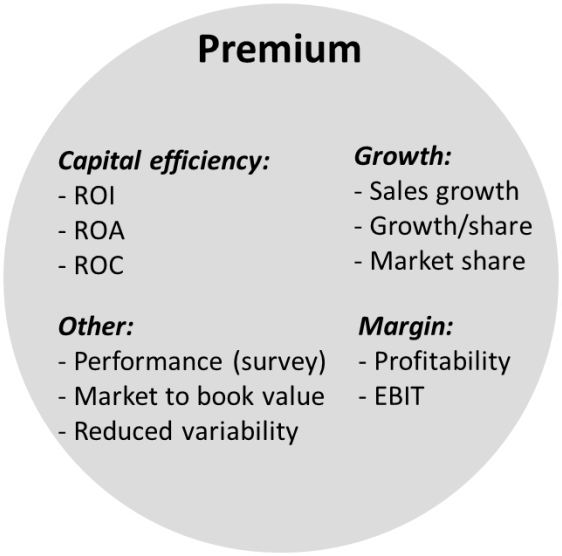

Figure 14: Literature findings premium mapped onto the coherence premium framework.

3.6 Conclusion

The mapping of literature findings onto the coherence premium concept is now finished. Apart from the structural contingency variables and several environmental factors it was possible to map the elements that were found. The combined result, together with the premium measures, is shown in Figure 15.

The core capabilities system box is filled with a few specific findings, some general considerations about core capabilities and several classification methods. The back-up from literature for the element therefore is limited. However, considering that the research field exists about two decades, RBV oriented literature findings make it clear that capabilities and resources are important aspects. The way to play has received very much attention in strategy research. Therefore this element is best substantiated by literature. Due to practical reasons, thus far the specificity of the way to play has not been implemented in research. Product and service portfolio has not been researched in a coherence related context. However part of the aspects fall under elements often researched by IO and contingency oriented scholars. The concept of diversification, which was added to product and service portfolio, has received a lot of attention. Although the links with the product and service portfolio are

the most indirect, it can be concluded that prior research indicates that it is an important aspect in the coherence context. Thus, all the elements received sufficient support to keep them in the model. However, the application of the coherence theory model is a clear and important limitation. The model will be reflected on in Chapter 8.

There is a large gap between the general coherence premium framework and the often specific literature findings mapped onto it. The usability of this literature review for the coherence measure to be developed therefore is limited. The fact that it was not possible to find a large amount of literature findings on this topic says something about the research field. It is clear that evidence for a relationship between fit (coherence) and firm performance (premium) has not yet been found. This also explains why so few indications were found for the reasoning behind the influence of factors affecting premium. This once more indicates that the strategy research field can learn from a new concept like the coherence premium, which can unify many specific findings with each other.

The next step in this research project, taken in the next chapter, is to extend and operationalize the coherence premium framework to make it widely applicable, but testable at the same time.

Figure 15 Relevant literature findings mapped onto the coherence premium framework.

4. Research Design and Methodology

In this chapter the research design adopted to test for the coherence premium in the large-scale sample is developed and discussed. It explains the approach to measure coherence with textual analysis. First the coherence model is operationalized using the insights obtained in the literature review. In Section 4.2 the scope choice for textual analysis, the data to be analysed and the data acquisition are explained. The analyses that will be performed to test the coherence premium relationship in the large sample are presented in the last section.

4.1 Operationalized model

Combining the Leinwand and Mainardi (2010, 2011) notion of coherence with the elements found in literature, the final operationalized model will be derived now. Subsequently the elements in the framework are discussed and operationalized individually.

4.1.1 Final Coherence Premium Model

The final model, shown in Figure 16, includes the measurement of coherence and control variables. Coherence will be measured by the three elements it consists of, core capabilities system, way to play and product and service portfolio. Since coherence cannot thrive when one of the three coherence elements is weak, even if the others are very strong, the coherence metric is operationalized as a multiplication of core capabilities system, way to play and product and service portfolio. To look at the coherence premium relationship, it should be corrected for influences by other factors. These factors could include year, size, diversification, industry, age, structure etc. depending on the analysis performed and the data available. The other relevant factors that will be used are given in Figure 16.

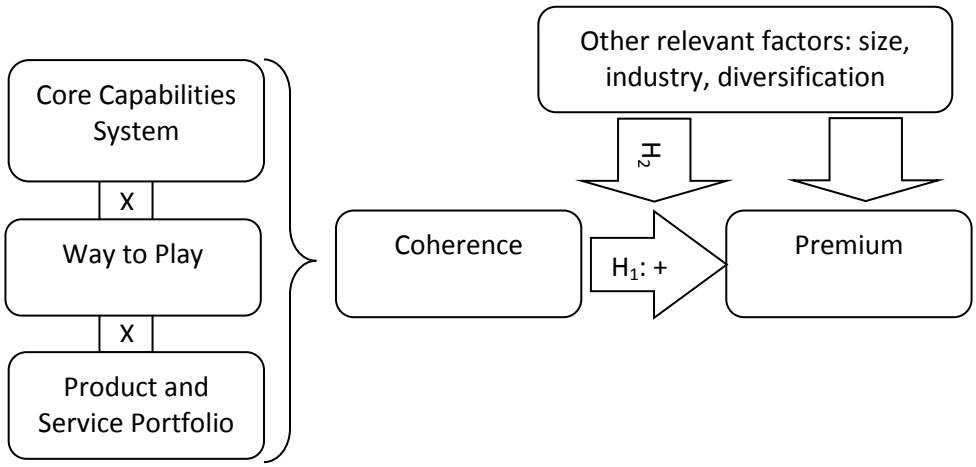

Figure 16 Operationalized coherence premium model with moderating influence of other relevant factors.

The main hypothesis is that coherence is positively related to premium. For the other relevant factors it is hypothesized that they can moderate this relationship. This leads to the following hypotheses:

H_1: Coherence is positively related to premium.
H_{2a}: Firm size moderates the relationship between coherence and premium.
H_{2b}: Diversification moderates the relationship between coherence and premium .
H_{2c}: The industry sector in which a firm is active moderates the relationship between coherence and premium.

The direct influence of the other relevant factors on (the average level of) the premium measures is not (directly) related to coherence and is thus excluded from the research goals. However, in the results and conclusion sections these effects, if present, will be presented as side findings.

4.1.2 Coherence Elements Explained

Coherence is the alignment of a firm's core capabilities system, way to play and product and service portfolio. Now the meaning of a core capabilities system, way to play and product and service portfolio should be defined. Additionally there remain other influences that have to be taken into account and premium needs more explanation as well.

Core Capabilities System

The integrated set of overarching core capabilities[9] will be used to look at core capabilities on an aggregate level. The classification at the first level is repeated here:

- *Supply chain management*
- *Marketing*
- *Portfolio management*
- *Manufacturing*
- *Selling*
- *Financial management*

Way to Play

A way to play can be found by combining several so-called strategic *puretones*[10]. These are highly specific ways to play. A simplification can be found in limiting the ways to play to these fifteen puretones.

Other alternatives can be found in the strategy literature. The Miles and Snow (1978) typology is widespread and classifies business units into four groups that exhibit different sets of strategic activities: Prospectors, Analyzers, Defenders and Reactors. The Miles and Snow archetypes have been used frequently in the past decades. This enables comparison with other literature studies.

A short description of the four Miles and Snow groups is given here; puretones related to the categories are indicated as well:

- **Prospectors** lead industry change, principally by launching new products and identifying new marketplace opportunities. Puretones include aggregator, customizer, innovator, risk absorber and solutions provider/integrator
- **Defenders** find and seek to maintain a secure niche in a stable product area. They stay within a limited product range, focusing more on resource efficiency and process improvements that cut manufacturing costs. Puretones include category leader, consolidator, experience provider, platform provider, premium player, reputation player, regulation navigator and value player/low-cost producer
- **Analysers** act as both prospectors and defenders, by both defending positions and sometimes moving quickly to follow opportunities. They are more likely to follow a second-but-better strategy. Puretones include customizer, disintermediator, fast follower and experience provider

[9] Developed and discussed in Section 3.2.2 Table 11
[10] Puretones are shown in Table 3

- **Reactors** are businesses that lack a consistent strategy, and respond to environmental influences as they arise. They are typically outperformed by the three strategic types with consistent strategies.

A simplified differentiation based on these strategic types could be between prospector, defender and analyser/mixed.

Another classification could be based on Porter's Generic strategies, which include differentiation, cost leadership and focus strategy, or the Treacy and Wiersema model, which differentiates between the strategies of customer intimacy, operational excellence and product leadership. The options presented here are by no means exhaustive.

Product and Service Portfolio

Fairly straightforward, the product and service portfolio is the package of different products and services a company sells.

This is closely related to the concept of diversification, which can be measured by the number of segments a company competes in. The simplest classification is between a narrow and a broad portfolio, or between diversified and undiversified firms. This can readily be assessed by looking at the number of SIC-codes a firm's activity spans. However the coherence premium concept does not differentiate between diversified and undiversified firms. Instead it is about applying the same core capabilities to all products and services offered. This cannot be measured with commonly available data.

Premium

Measurement of premium, or above normal financial performance, is relatively straightforward as there are many metrics readily available. For a good comparison between different companies it is important that the measures are relative and not absolute. In the literature chapter it was found that metrics used often include growth, profitability/margin and efficiency measures. Another conclusion drawn there was that the use of multiple premium measures was beneficial to improve the validity of conclusions and to decrease biases. Therefore, the choice was made to include at least a growth, a profitability/margin and an efficiency measure. For the profitability measure EBIT margin was chosen, because it enables validation of the results obtained by Leinwand and Mainardi (2010). The most frequently used growth measure is sales growth. Sales growth and EBIT margin fundamentally drive profits (Varaiya, Kerin and Weeks 1987). For the efficiency measure, return on assets was chosen, as it is less dependent on financial structure and leverage. Another premium measure, P/E ratio, which reflects a firm's valuation by the stock market, was added out of interest and curiosity. All of the premium measures are mentioned by Young and Basarab (2009) in their Business Acumen training program. The premium measures will be compared amongst a peer group, consisting of companies in the consumer and retail industries for the full large sample.

These premium measures will be briefly explained here. The EBIT margin and sales growth are the most important metrics as they are the fundamental value drivers (Koller et al. 2010). They are fundamental because sales x margin on these sales is what really drives returns. Return on assets measures the efficiency with which assets are used. Higher efficiency is an indicator of better performance, but there is no causality implied. The P/E ratio is a measure for the valuation by the stock market of a company relative to its earnings. Generally speaking, a company with a higher P/E ratio than its direct competitors is performing better than its peers (direct competitors), because investors are willing to pay a higher multiple of its yearly profits. However, the P/E ratio can be influenced by many other factors such as the

future outlook for the company. Additionally, it is not a measure of the current performance only. Performance in previous years and future outlook are taken into account as well.

Earnings Before Interest and Taxes margin (EBIT margin)

$$EBIT\ margin = \frac{ebit}{sales}$$

The EBIT margin is one of the two fundamental value drivers. It measures the earnings before interest and taxes relative to sales. The EBIT margin is dependent on the industry a company competes in. There are significant differences between the margins on for instance chemical commodities and clothing (Young and Basarab 2009).

Sales Growth

$$sales\ growth = \frac{sales\ current\ year}{sales\ previous\ year} - 1$$

The sales growth is the other fundamental value driver. It is the most common measure for company growth. Growth is important for a company to keep improving its profits. Yearly growth does not fundamentally differ between industries. However in some industries it might be easier to grow than in others due to, for instance, the competitive intensity. Growth is different for companies of different sizes; whereas a start-up might double its sales in a year, the Coca-Cola Company will certainly not be able to achieve such growth. With increasing size the growth percentage generally decreases.

Return on Assets (ROA)

$$return\ on\ assets = \frac{net\ income}{total\ assets}$$

Return on assets measures the efficiency of the use of assets. Between industries the figures differ widely. Capital intensive companies like manufacturers have significantly lower returns on their assets than companies that are more labour intensive, such as advisory or design firms (Young and Basarab 2009). Another drawback is that ROA uses the book value of total assets. Older firms, with depreciated assets, will therefore obtain higher ROA values.

Price/Earnings Ratio (P/E ratio)

$$P/E\ ratio = \frac{common\ shares\ outstanding * share\ price}{net\ income}$$

The P/E ratio is a measure for the valuation by the stock market of a company relative to its earnings. It incorporates the results of a company in the past few years and the investor's interpretation of the future outlook of this company. This means that the P/E ratio does not reflect current performance alone; past and anticipated future performance have a large influence as well. Additionally the P/E ratio is dependent on capital structure. At high levels of leverage or when profits decline substantially the P/E ratio increases. Therefore it is more difficult to make comparisons based on this metric than on the others mentioned in this section. Still, the P/E ratio is used as an additional premium measure, because it is used frequently throughout the financial world.

Other relevant factors

As became apparent in the literature review, there are aspects that do not fit well into the coherence framework. Some of them can be taken into account. These include fiscal year, size and industry. Before taking them into account, it must be ensured that it is sensible to do so. A clear differentiation must be made between direct effects of other relevant factors, which will not be tested in this study, and indirect effects that moderate the relationship between coherence and premium. As there is no information available on this moderating behaviour yet, the possibility of the effects is briefly discussed. It is not means to develop hypotheses, merely to assess whether it makes sense to expect a moderating effect.

It does make sense to differentiate between industries. Perhaps there are certain industries in which coherence is more critical and/or more valued. Regarding size, it could be that coherence becomes more important at larger firms, because the distances within the company are larger and focus can be added by coherence. On the other hand, coherence could be even more important for smaller firms to be profitable. Diversification could pose a challenge, because being coherent is likely to become harder with an increasing number of products and services.
Additionally it could be the case that coherence is stronger or weaker related to a subset of the premium measures for certain industries, size groups or a diversification group.
For fiscal year it is harder to come up with arguments; it remains unclear why the year would have any influence on the relationship between coherence and premium.

Based on the aforementioned, it was decided to compensate the results obtained with the full large sample for different size, industry and diversification groups.

4.2 Textual Analysis

Surveys were used to substantiate the hypotheses in the majority of articles published in the strategy research field. Therefore investigation of the 'coherence premium' on a quantitative basis is a complementary approach; it will add value to existing research. Textual analysis is the research method used in this research project. It is a promising alternative for surveys, because of the following advantages:

- Bias is different from surveys, making the methods complementary (Jauch et al. 1980).
- No dependency on response rates.
- Potential method to add to and change the research methodology palette (Jauch et al. 1980, Bowman 1984).

Furthermore it enables quantitative analysis by large-scale application of textual analysis.

Textual analysis is based on counting words. When searching through a publication with a set of keywords typical for coherence (elements), the number of hits on this set of keywords can be used to indicate the emphasis put on coherence (elements). This coherence emphasis for the companies in an industry based peer group will be compared with their premium to assess the effects coherence has on premium.

Looking at companies can be done with an outsider perspective or with an insider perspective. The outsider perspective analyses external data, e.g. press releases, annual reports, external databases, whereas an inside perspective can be taken by performing case studies, interviews or surveys. The aim of this research project is to test for the existence of a coherence premium with a large-scale dataset. Case studies and interviews were discarded for this reason and because they were used in Leinwand and

Mainardi (2010). Another option, performing a survey, is less appealing, because they are used very often in strategy research and thus not new or innovative. Additionally, they have to be filled out by people, which limits the speed of data acquisition as well as the amount of data that can be acquired. Furthermore Booz & Company has recently developed a survey called the 'Coherence Profiler' already. Alternatively, taking an outside perspective, numerical and textual analysis can be performed using for example analyst reports, newspaper articles, annual reports or financial filings. Numerical analysis is done professionally by analysts already and can tell us little about coherence.
Textual analysis is a promising approach. Content analysis (of annual reports) can be very useful for understanding some issues of corporate strategy as primary or supplementary information source (Bowman 1984). Besides, textual analysis can be done with large datasets with computer scripts. This enables quantitative analyses with statistically significant conclusions. Moreover, it is a new method that, as far as we know, has not been used in a related context. On top of that, it can add to the methodology portfolio available for strategy research.

All in all we see potential for using textual analysis on company publications to measure coherence. The assumption for using textual analysis is that companies disclose sufficient indications about their own coherence in their publications. Although they were not related to coherence, scholars have used textual analysis year reports (Bowman 1984) and other publications (Jauch et al. 1980). This indicates that these texts contain enough information to perform research with.

The nest steps, selection of a data source and the data acquisition, will be explained in the next sections.

4.2.1 10-K Item 7: Management Discussion and Analysis

After deciding to use textual analysis the type of data to use in the textual analysis has to be chosen. Newspaper articles, press releases, annual reports (used by Bowman 1984), U.S. Securities and Exchange Commission (SEC) filings and analyst reports are considered here, because initial screening shows that elements relevant to coherence are discussed. Therefore, they could provide sufficient indication of how relevant and important these coherence elements are to the company.

- Newspaper articles usually contain outsider information, written by a wide range between non-experts and experts that are generally not (positively) biased to the firm. Due to this range in expertise the newspapers and/or authors from which/whom to include articles should be limited.
- The CEO-letter is a two to four page item in the annual report, which gives a summary of the past year's performance and reveals the plans of the CEO/company for the coming year. The CEO is a major stakeholder in the firm and he gives us an expert's insider view. As this is an official release, it can be safely assumed that disclosed information is factually correct. However, the CEO-letter will still be written to come forward as positive as possible; it has a large marketing bias. This may include omission of negative aspects and extensive attention paid to positive elements.
- The Management Discussion and Analysis (MD&A) SEC 10-K filings is a rich (¬20 pages) and relevant (management view on strategy) source of information. Content-wise it is similar to the CEO-letter, but it contains much more information. Since it is a large item in an official filing the positive bias will be less evident than in the CEO-letter.
- Analyst reports, give an outside-in view from an industry expert. These tend to be more focused on the financial aspects of the firm. Usually the analyst has close contact with the company he/she follows, which suggests a positive bias towards the company.

Considering the options mentioned above, the MD&A will be used as data source for this research, because the potential value of the content is much larger than that of the other sources. No other source gives such a rich insider perspective on the firm. The MD&A gives an extensive inside expert perspective on the entire firm, which is very valuable. As in all company publications, the MD&A will put the company in a positive light where possible. Because positive or negative words are not explicitly measured, the effect on coherence appears to be very limited.

4.2.2 Acquiring the Data

For our large-scale aspirations it is crucial to be able to obtain a large amount of valid data. The large dataset was retrieved from the U.S. Securities and Exchange Commission (SEC). It is stated on their website that 'All companies, foreign and domestic, are required to file registration statements, periodic reports, and other forms electronically through EDGAR. As they state it: *'Anyone can access and download this information for free. Here you'll find links to a complete list of filings available through EDGAR and instructions for searching the EDGAR database.'*

The SEC filings are publicly accessible and with a script written in programming language Perl, see Appendix IV.A, exactly 100,665 K-10 files were downloaded from the Electronic Data-Gathering, Analysis and Retrieval system (EDGAR) database between June 29 and July 3 2012. As the 10-K filings also include financial reporting, the common financial performance metrics are available from the EDGAR database or can be calculated with the data it has available. Section 6.1 will discuss the data in more detail.

The batch download was just the first step. The next step is to correctly and accurately extract the MD&A's from the 10-K forms. The extraction of the MD&A and the counting of the words in the MD&A were done in a single Perl script[11]. The extraction of the MD&A proved to be quite difficult. Working together with Joost Impink, who wrote the scripts, many iterations were required to perfect the method. In the end the script was able to handle all sorts of separations between 'Item' & '7' and 'Item' & '8' to signal the start and end of the MD&A. Additionally, occurrences in conjunction with dequalifiers like 'see' and 'continued' were neglected and the minimum length for the MD&A was set to 10,000 characters[11].

Two test sets were used to assess the script: the first is the set of companies from the 'The Coherence Premium' article from 2003-2007. The second is a random sample of 100 10-K's. The script was able to successfully extract the existing MD&A's for the test set except for one. For Coca Cola 2007 some financial data and several paragraphs of item 6 were mistaken as belonging to the MD&A. The extra text extracted was less than 5% of the original MD&A and contained mostly numbers. This minor mistake in one of the 36 data points did not affect the results from the analyses with the first test set. The results for the second test set are better: 98 of the 100 10-K filings were successfully processed, judged by file size comparison to correctly extracted MD&A's on three significant digits. 85 MD&A's were successfully extracted; this includes correct failures due to MD&A's by reference. Two MD&A's were not extracted, while they were present. Thirteen 10-K filings were 10-K SB filings, which could not be extracted, as will be explained shortly. Most importantly, no errors were made in the extraction process, which indicates that the reliability of the extracted MD&A's is high. Final result is 66 MD&A's as nine 10-K filings had no MD&A. Brown and Tucker (2011) were able to extract MD&A's from 73% of the 10-K filings. They state that in most of the cases in which the MD&A's could not be extracted the 10-K filing did not have an MD&A. In the end the script enabled the extraction of 6,974 MD&A's from 10,328 10-K's (67.5%). The analysis was limited to this subset due to the long run time, which increased almost exponentially with

[11] A more detailed description of the script and its functions can be found in Appendix IV.B.

the 10-K file size. The subset consists of companies active in the consumer and retail industry, just like the companies in the Leinwand and Mainardi (2010) article, which is also used as a test set in this thesis. The SEC uses different industry segmentations, the consumer and retail is translated into the FamaFrench-48 industries food, beer, retail, meals, soda, clothes, textile and paper.

These results seem to be in line with Brown and Tucker (2011). However, upon closer examination it was found that there exist several types of 10-K. The following types can be found in the original 100,665 item dataset:

- 10-K (72.03%)
- 10-K 405 (12.64%)
- 10-K SB (13.55%)
- 10-K SB 405 (1.65%)
- 10-K T (0.09%)
- 10-K T 405 (0.03%)

The suffix 405 indicates that one or more officers or directors of the public company failed to disclose their insider trading activities within the required time period. Content-wise there is no difference with the 10-K form. SB denotes small business and does affect the 10-K content. For small businesses some of the 10-K items are not required. They do contain an MD&A, but this has become item 6 instead of item 7. This means that MD&A of these 10-K SB filings cannot be extracted with the script. The suffix T indicates a transition, meaning a change in fiscal year-end. It is submitted instead of or in addition to a normal 10-K form. When the transition is complete the company returns to filing regular 10-K forms (Securities and Exchange Commission).

The transition 10-K forms are different from normal 10-K forms and occur very infrequently. Therefore they are excluded from our analysis. This decreases the dataset with fourteen items to 6,960 items. The 10-K SB forms are not present in the data, because the MD&A cannot be extracted. Assuming similar 10-K type frequencies in our subset, this means that roughly 83% of the 10-K and 10-K 405 forms have an MD&A. This is higher than the 73% reported by Brown and Tucker (2011).

4.3 Analysis Method: ANOVA

In this section the analysis methods will be explained briefly. The Analysis of Variance (ANOVA) will be discussed briefly. In the results section the ANOVA is used to look for significant premium differences between the high and low coherence groups in the large dataset. Subsequently it is used to look into other factors that might cause these differences. The effect of coherence will be tested in different size, industry and diversification groups. Finally, several case studies will be performed, which take into account all these factors simultaneously.

The ANOVA is a method which tests whether the means of two or more groups significantly differ from each other. More specifically it tests the null hypothesis of equal means: $H_0: \mu_1 = \mu_2 = \cdots = \mu_n$

The ANOVA assumes that (Laerd Statistics):

- The independent variable consists of categorical independent groups
- The dependent variables are interval or ratio scales
- The dependent variables are approximately normally distributed for each categorical group of the independent variable
- Variances between the independent groups are equal
- Cases are independent

The data certainly fulfils the requirements of assumptions one, two and five. Apart from the P/E ratio, the dependent variables appear to fulfil assumption three as well. The equal variances assumption can be checked during the ANOVA with Levene's Test of Homogeneity of Variances. If the equal variances assumption is not met a nonparametric alternative for the ANOVA is the Kruskal-Wallis test. This test follows the same procedures as the regular ANOVA but does not assume any type of distribution or equal variances. Unfortunately the Kruskal-Wallis test does not give any information beyond the significance.

In the ANOVA, three SPSS tables are of interest:

- First of all, the ANOVA table. With every dependent variable it gives us the sum of squares, df, Mean Square, F and Sig. between groups, within groups and total. The important column is the Sig. column. If the significance value in the Sig. column is smaller than .05 the null hypothesis can be rejected with 95% certainty. This means that the difference between the means of the two groups is significant.
- The ANOVA table does not contain more information about these means. The Descriptives table is useful here; it gives the sample size (N), mean, standard deviation, standard error, 95% confidence lower and upper bound of the mean, minimum and maximum value for each of the independent variable groups and the total. The mean column is most important, since it tells us what the means are of the groups that had significant differences (or not).
- Last there is the Levene's Test of Homogeneity of Variances Table. If the significance value in the Sig. column is larger than .05 the null hypothesis, which is homogeneity of variances, cannot be rejected with 95% certainty. This means that the variances between the independent groups are equal and that assumption four has been met.

A one-way ANOVA is the simplest ANOVA; it compares the means between one grouped independent variable. In our case the coherence score is divided into three equally sized groups consisting of low, medium and high coherence scores. The one-way ANOVA will be used to compare the means of the low and high coherence groups.

4.4 Conclusion

In this chapter Leinwand and Mainardi (2010) and the relevant literature from Chapter 3 were used to operationalize the coherence theory model. Coherence was conceptualized by the multiplication of the three coherence elements. In this research these coherence elements are represented by the attention paid to them in the MD&A of the 10-K filing. It is measured by the number of hits obtained for general keywords that are indicative for emphasis on this coherence element. Finally, the method of analysis, analysis of variance (ANOVA), was explained.

In the next chapter a coherence measure will be developed, using the operationalized model and textual analysis.

5. Developing a Coherence Measure

In this chapter the development of the coherence measure is explained. It answers the question how the coherence premium can be measured. The first section discusses the use and selection of keywords to obtain a set of coherence indicative keywords. The second section takes into account a large number of methodological options. It decides on the final measurement construct by testing the options identified earlier with a test set from Leinwand and Mainardi (2010). This test set is described in more detail in Appendix II.

5.1 Selecting Keywords

As a means of measuring the relative importance of coherence to a company by, the method of counting words that are indicative for coherence or for its underlying elements, was chosen. It is 'relative' importance, because the MD&A length and an industry peer group for comparison are taken into account. The full sample consists of firms from eight FamaFrech-48 industries in the consumer and retail industry, which is the initial peer group for comparison. For this measurement a reliable construct has to be developed. In this process it is important to ensure that the words included are as comprehensive as possible, truly reflect coherence and are not used in another context. This was done by first creating a longlist of keywords, combining Leinwand and Mainardi (2010, 2011) with cues from the literature research, including synonyms and personal additions. This keyword longlist was subsequently manually tested and transformed into the final keyword list. This section will explain in more detail how this was done.

5.1.1 Developing a Keyword Longlist

The discussion of the coherence elements in Section 4.1.2 was used to create the longlist for coherence and the underlying elements. Leinwand and Mainardi (2010, 2011a) and other literature sources were used to find appropriate keywords. Subsequently the list was expanded by incorporating synonyms. This resulting keyword longlist was divided into several categories. These included general coherence stating symptom keywords, general keywords for the three coherence elements and specific keywords for several categories underlying the way to play and core capabilities system elements. For the underlying categories of the core capabilities system the first level classification derived in the literature chapter, see Table 11, were used: marketing, supply chain, manufacturing, sales, portfolio management and financial management. The goal is to measure the strength of the specific core capabilities system of the firms with these categories. Specific way to play categories include Miles and Snow defender, analyser and prospector, Leinwand and Mainardi's puretones, Porter's generic strategies, the strategy dimensions of Venkatraman (1989a) and the Treacy and Wiersema (1993) model.

The longlist was further expanded with information from annual report CEO-letters. The CEO-letter was used for this purpose, because content-wise it is a summary of the MD&A. Its smaller size allows for a larger set of different company-year combinations to be included within a reasonable amount of time. First six CEO-letters in annual reports of three companies[12] were scanned for sentences indicating a way to play or capability and the keywords from these sentences were added to the keyword longlist. Two additional CEO-letters, TJX (2006) and Colgate-Palmolive (2009), yielded seven and three new keywords respectively. As this amounts to only 1-2 % of the total keyword list it was concluded that the region of diminishing returns was reached and that the keyword list was thus large enough to identify the majority of indicative keywords in the MD&A. The longlist can be found in Appendix III.A. It incorporates 484 unique keywords in total, with some keywords present in more than one group.

[12] Wal-Mart: 2003, 2005; Coca-Cola: 2005, 2006; Wrigley: 2003, 2007

5.1.2 Improving the Keyword Longlist

With the longlist as a starting set the annual reports of Coca-Cola (2004), Con-Agra (2007), Wal-Mart (2011), TJX (2006) and Colgate-Palmolive (2009) were subsequently coded manually. In the CEO-letters of these annual reports a search was performed for all keywords and every hit was coded manually. This manual coding comprised of evaluating the context (sentence) surrounding a keyword and differentiating between a positive and negative (unrelated, false positive) association of this keyword with the coherence element it should indicate. On top of that a rating of 1-3 was given to indicate the strength of the statement, ranging from vague to clear and convincing. All the coding was done by the author personally.

The next step was to clean this keyword longlist. The two measures used in the coding exercise were used to clean the keyword list. The first is the percentage of non-indicative hits or false positives; the second is the average score of correct keyword hits (false positives excluded). Combinations of correct keyword percentages of 50, 70 and 90% and score thresholds of 1.5, 2 and 2.5 were used. The extreme combinations of the highest score threshold with the highest false positive percentage and vice versa were discarded. The effect of the cut-off score on the size of the keyword shortlist was found to be much larger than the effect of the applicability percentage.

The cut-off values and incorporation of general or specific words do not stand alone. They constitute two of the many choices that will be considered and evaluated in the next section using the test set.

5.2 Computing a Coherence Score

Now that a longlist of keywords indicative for coherence has been developed, a coherence score measure must be developed with this basis. There are many methodological options, from which the most appropriate ones should be selected. To make the definite choice for a coherence measure, the test set will be used to test the performance of different sets of options[13]. Most of the options give us a choice between using less, more indicative keywords or more, less indicative keywords. These need to be balanced, to get an accurate and reliable measure.

5.2.1 Choices to Make

In the decision which keywords to take into account there are four choices to be made:

- In the keyword list used in the coding procedure, all three coherence elements were present with general words. Both core capabilities system and way to play were present with keywords in specific subcategories, indicative for certain capabilities or ways to play. The first choice is whether to take into account the general keywords only, the specific keywords only, or both.
- The coding is a task that has been performed on a single keyword basis. As this leads to a small number of hits per keyword, one could argue that a more aggregated approach would yield more hits per keyword group and thus give more confidence in the rating of words. To give an example, the rating of collaborate, collaborates, collaboration, and collaborated could be combined. The second choice is between single, aggregated and aggregated plus. Aggregated plus is aggregated with some manual improvements: familiar keywords that are usually used in a different context separated here.
- The third choice is whether to include keywords that receive no hits or to require at least one or two hits for a keyword to be included. Higher requirements further improve the reliability, but decrease the length of the final keyword list and therewith limit the breadth of the final coherence score measurement.

[13] More information regarding the test set can be found in Appendix II.

- Last, there is the cut-off value. As the keyword hits are rated on applicability (0-1, irrelevant-relevant) and strength (1-3, vague-convincing), there are two measures on which to (de)select a keyword. As a minimum to be indicative, the keywords should be applicable in 50% of the cases. Additionally a score of 1.5 on average, for the applicable hits only, should be obtained. Using applicability constraints of 50%, 70% and 90% plus average score constraints of 1.5, 2 and 2.5 there are nine options. However, the extreme combinations of 50% plus 2.5 and 90% plus 1.5 will not be taken into account as it does not make sense to combine the most strict and the most tolerant options of two parameters with similar importance.

In the part where a score is computed from the keyword hits:

- A first choice will be made between the coherence score construct, which multiplies the coherence elemental scores with each other, and the simple counting and adding of the hits obtained with the keyword list. Addition of the three coherence elemental scores will be tested as well. These will first be normalised to equalize their impact. The operationalization of the coherence score construct in the coherence framework will thus be reconsidered if necessary.
- A comparison will be made between normal counting and summing the 10logarithm of the hits for every individual keyword for a coherence element. The latter will decrease the dependency of the score on one or a few keywords that occur (very) frequently.
- The third choice is between comparing the coherence score to the premium measures in that same year, the year after or two years later. This is done to take into account the time it might take to achieve coherence, after discussing and analysing the company performance in a coherent way. Another alternative here is to take the average coherence score and average performance metrics over a time period of several years. Unfortunately, the large dataset was not very suitable for this approach.

Altogether there are over 3000 different combinations possible. There is no simple method to compare such a number of options simultaneously. Therefore a base case was constructed. The options for six of the seven choices were kept constant in this base case, while varying the options for the seventh choice to evaluate the effects of its options. In this manner all seven choices were evaluated using this base case.

The base case incorporated the following options:

- General words only
- Include keywords with one or more hits
- Aggregated form of keywords
- Cut-off at 70% and average score of two
- Coherence construct from general model (multiplied elements)
- Normal counting per keyword
- Compare coherence and results from the same year

5.2.2 Choices Made

Options were evaluated[14] based on the predictive value (R^2) of the linear regression lines. The correlation of coherence with premium, or the slope of the regression line, was not taken into account, because the goal was merely to find a coherence metric that fits well with the data and not the one which shows the largest coherence effect[15]. The base case is shown in Figure 17 with the ROA, EBIT margin and sales growth metrics[16].

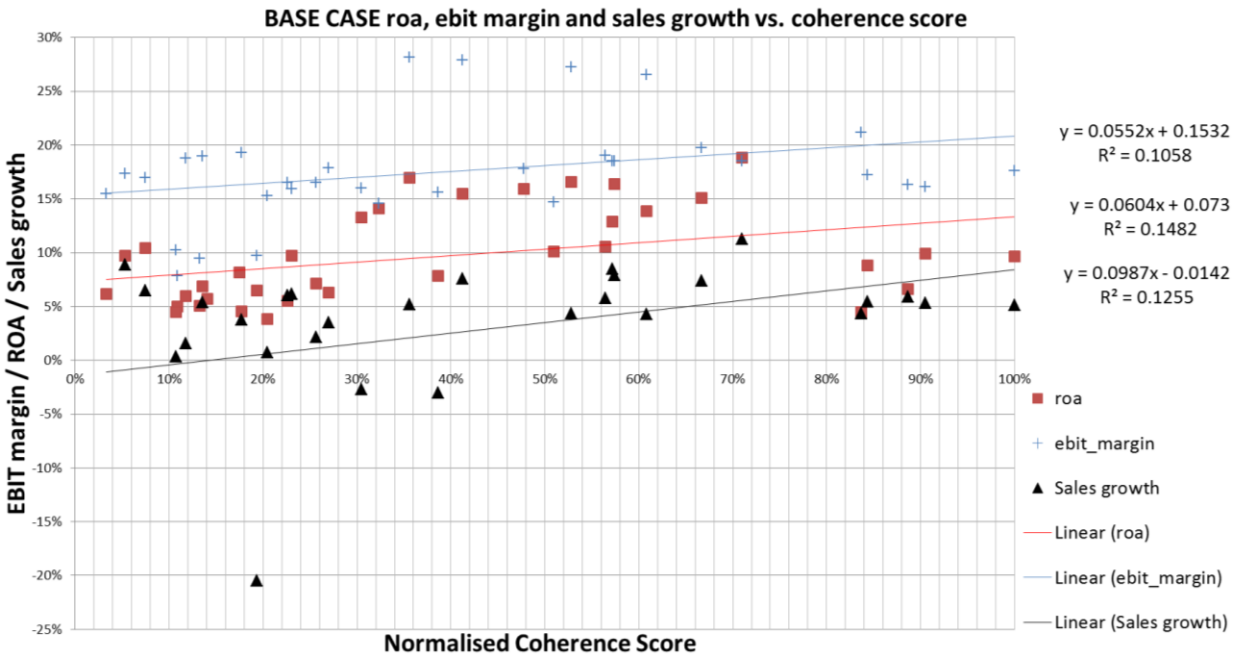

Figure 17 Base Case coherence score versus ROA, EBIT margin and sales growth.

Table 15 R^2 values for different variations on the base case and for the improved case.

	P/E ratio R^2	EBIT margin R^2	ROA R^2	Sales growth R^2
Aggregated plus form of keywords	0.1138	0.1273	0.1613	0.1214
Cut-off at 70% and average score of 2.5	0.2195	0.2316	0.2984	0.0835
Logarithmic counting per keyword	0.1593	0.2366	0.2466	0.1133
Base Case	0.0889	0.1058	0.1482	0.1255
Improved Case	0.1590	0.2481	0.2178	0.0961

The options that resulted in improved R^2 values compared to the base case are stated in Table 15. It stands out that the R^2 of sales growth is relatively low and did not improve much upon varying the options. This can be attributed to the large negative value, which decreases the fit considerably. However, it cannot be regarded as an outlier, because the sales growth value is within reason and the information is accurate. The company in question, ConAgra, really lost over 20% of its sales in 2005.

With these improvements, the *improved case* was constructed; it is shown in Figure 18.

[14] All the cases were evaluated with Microsoft Office Excel 2010.
[15] More results can be found in Appendix II
[16] P/E ratio is excluded because the different y-axis requires an additional graph which shows the same trend.

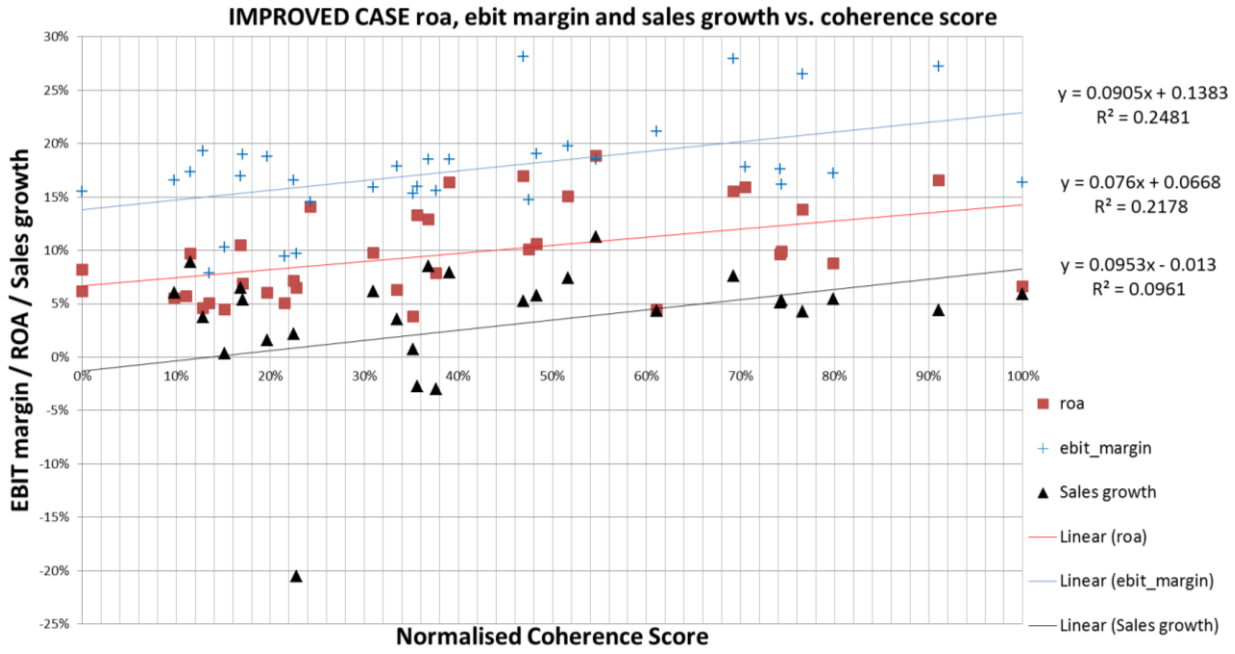

Figure 18 Improved Case coherence score versus ROA, EBIT margin and sales growth.

The improved case incorporated the following choices:

- General words only
- Include keywords with one or more hits
- Aggregated plus form of keywords
- Cut-off at 70% correctness and an average score of two
- Coherence construct from general model (multiplied elements)
- Logarithmic counting per keyword
- Compare coherence and results from the same year

This means that the options that improved the predictive value were all implemented, except for the cut-off values. Although the cut-off value at an average score of 2.5 significantly improved the predictive strength it also severely distorted the data. Seven of the thirty-seven data points had a coherence score of zero and on top of that the data shows a large positive skew. The zero scores are attributed to the quickly decreasing number of keywords with increasing cut-off score requirements. A lower number of keywords improve the chance of obtaining a zero score. Such distortion of data can spoil statistical analyses. Therefore the cut-off was kept at an average score of two. These choices left 134 keywords in total for the three coherence elements.

From table 15 it can be seen that logarithmic counting per keyword scores a little better than the improved case for the test set. The only difference between the two is that for the improved case the aggregated keywords were manually checked and separated several times if the grouping was inappropriate considering the normal use of these words. Because this is an improvement and the differences are small it was decided to use the aggregated plus form and thus the improved case.

To ensure the (local) optimum was reached, the base case procedure was repeated with the improved case as base case for the choices that were not very clear in the base case: cut-off values, inclusion of keywords with at least zero, one or two hits, normal/logarithmic/square root counting per keyword, comparison of coherence and premium in the same year. No further improvements were made. The final keyword list is shown in Appendix III.B.

5.2.3 Final Coherence Score Calculation

With the choices made, the final coherence score calculation can be explained. This is best done by giving an example.

> Consider company A, which gets 80 (8*10 hits), 120 (12*10 hits) and 230 (13*10 hits, 1*100 hits) keyword hits in its 10,000 word MD&A on core capabilities system, product and service portfolio and way to play respectively. Taking the logarithm of the number of hits for individual keywords scores of eight (8*10log(10)), twelve (12*10log(10)) and fifteen (13*10log(10)+1*10log(100)) respectively are obtained. Its coherence score then is $\frac{8}{10{,}000} \times \frac{12}{10{,}000} \times \frac{15}{10{,}000}$. Suppose the maximum coherence score in the peer group is $\frac{16}{20{,}000} \times \frac{24}{20{,}000} \times \frac{60}{20{,}000}$. Dividing the coherence score of company A by the maximum coherence score in the peer sample gives the normalised coherence score, 0.5 in this case.

This normalised coherence score is the final coherence metric. It will be used in all the analyses.

5.2.4 Automated Coherence Score Calculation

The Perl scripts used for counting the coherence keywords, written by Joost Impink, can be found in Appendix IV.C. The total keywords per MD&A and the score for the three coherence elements provided by the scripts were imported in MS Excel 2010[17] to calculate the coherence score. First the logarithm of keyword hits per keyword is calculated and added up per element. Then the elemental scores are divided by the total words in the MD&A. This division is performed to compensate for the length of the MD&A, as a longer MD&A is likely to contain more coherence element keywords. Subsequently the three elemental scores are multiplied to calculate the absolute coherence score. Finally the scores are normalised by dividing by the highest coherence score in the peer sample, consisting of consumer and retail companies.

5.3 Conclusion

In this chapter the construction of the coherence measurement tool was explained. In Chapter 4 it was decided to represent coherence by the attention paid to coherence elements in the MD&A of the 10-K filing. It is to be measured by the number of hits obtained for general keywords that are indicative for emphasis on this coherence element. This 'indicative value' of the keywords was measured by using 5 CEO-letters from annual reports as a proxy for the MD&A. Both the relevance (yes or no) and strength (1-3) of indication were measured.

[17] Microsoft Office Excel™ is a registered trademark of Microsoft Corporation. In this thesis the 2010 version was used.

Important choices that were made include:

- Requiring at least one hit in the CEO-letter measurement to be included in keyword list
- Only words indicating attention paid to elements in general were included, as opposed to attention paid to specific capabilities or ways to play.
- For the three coherence elements together, 134 words were left to be used in the analysis by applying a cut-off value of 70% relevance and an average score of two.
- Words are counted by taking the logarithm of the numer of hits per keyword.
- The sum of keyword scores for an element is that coherence element's score. After multiplication with the other elemental scores and division by the largest value obtained the normalized coherence score is calculated. This normalized coherence score is used in the analyses in Chapter 7.
- The coherence score will be compared to the performance in the same year

The next chapter will describe, analyze and clean the dataset; this is the last step towards the analyses.

6. Data Description

In this chapter the approach to data cleaning and transformation, which prepares the data for the analyses is explained in Section 6.1 and performed in Sections 6.1 through 6.5. Company-year combinations, coherence score data, premium measures and other relevant factors are analysed, cleaned and discussed. Before starting analyses, the data has to be reviewed to assess the accuracy, reliability and distribution of the data. This is important, because the data has to meet the assumptions for the statistical analyses. Furthermore outliers can have a large influence on the end results, something that should be prevented. The approach to handling these issues is presented in the first section here. SPSS Software[18] is used for the data analysis and cleaning, because it is known for its strength in statistical analyses and straightforward user interface.

6.1 Approach for Inspecting, Cleaning and Transforming Data

Before performing any tests data often has to be cleaned. Outliers can have a very strong influence on analysis results. However, they may not be disregarded that easily, because they can still contain valuable, accurate information. In the cleaning process the reliability and accuracy of data points is the key. As the dataset is large it is not problematic to exclude data with doubtful reliability. First the data will be investigated for possible causes of abnormal and unlikely values. Second, the distribution of the extreme values, taking into account the leverage and influence of single data points at the extremes, will be looked into. The distance from the mean, measured in the number of standard deviations, can be helpful here, even if the data does not show a (perfect) normal distribution. Furthermore in regression analysis outliers can be evaluated with the help of cook's distance (a measure for the influence of a data point), residual values and distance from the mean (several standard deviations)

Many analyses in SPSS have normality assumptions. If the data is not normally distributed this may cause the analysis to falsely conclude the null hypothesis should be rejected. Small deviations from normality are usually unproblematic, especially with large datasets. If data is not normally distributed it can be transformed, e.g. taking the square root, logarithm, inverse etc., to increase its normality. A potential downside is that this can change the interaction between the variables. Our dependent variables have a clear and well understood meaning, which will be lost upon transformation. Therefore it was chosen not to transform it. This means that analyses for non-normally distributed data could be necessary. Fortunately, SPSS has a nonparametric (assumes no specific distribution) alternative for most of its parametric analyses.

The data used for the analyses is now inspected and, if necessary, cleaned according to the presented approach. The information acquired, partially constructed by Joost Impink, for most of the 10-K SEC-filings in our large-scale dataset includes the following:

- company name
- fiscal year (1996-2010)
- SIC-code (4-digit)
- FamaFrench-12 industry
- FamaFrench-48 industry
- number of segments
- total assets (current and previous year)
- sales (current and previous year)
- common equity

[18] SPSS Software™ is a registered trademark of IBM Corp. In this thesis version 19 was used.

- net income (NI)
- common shares outstanding
- end-of-year share price
- earnings before interest and taxes (EBIT)

With this information and the computed coherence score coherence can be related to the following premium measures: EBIT margin, sales growth, return on assets and P/E ratio. The coherence and premium data will be discussed shortly. First the Company-year combinations are analysed and presented, followed by the coherence score data and EBIT margin, sales growth, return on assets and P/E ratio data. Last, the data concerning other relevant factors, size, diversification, industry and year, are discussed.

6.2 Company-Year Combinations

The dataset for which the MD&A's were extracted from the 10-K contains 6,960 company - fiscal year combinations. Examining the data 37 duplicate company – fiscal year combinations were found referring to different 10-K forms, which results in different coherence scores. As there is no means to assess which of the two or three is correct, these 37 duplicates were removed, reducing the dataset to 6,923 items.

There are no survival or occurrence restrictions to companies occurring in the large sample. This means that the number of years a company occurs in the data differs widely. This can be seen in Figure 19. In total there are 1251 companies, occurring 5.5 years on average. At first sight the data seems to be skewed towards companies that occur infrequently. However, this is untrue, since companies that appear for more years have more data points each. The data point origination by number of years the company has in the dataset is shown in Figure 20.

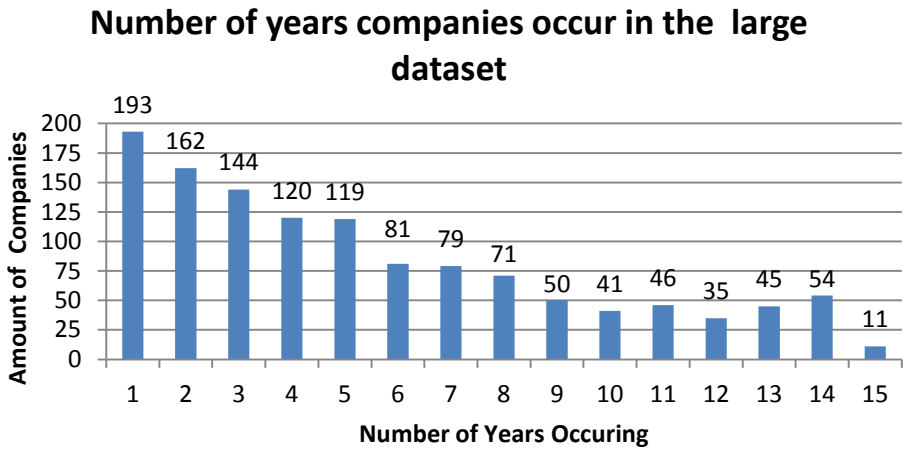

Figure 19 Histogram of the **number** of companies occurring in the data for different number of years.

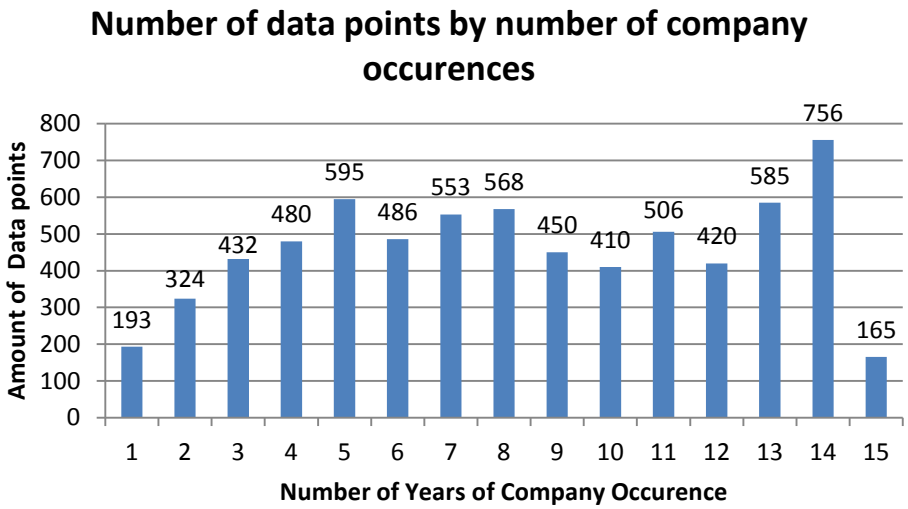

Figure 20 Histogram of **number** of data points for different numbers of company occurrences.

It should be noted that these company occurrences are not consecutive by definition. There are several reasons why most companies do not occur for fifteen (consecutive) years:

- The MD&A cannot be extracted successfully from all 10-K forms. The random 100 test set indicated that 2 MD&A's that were present were not successfully extracted; there will be at least two, likely more, 10-K filings that are not successfully extracted.
- Not all companies file their Management Discussion and Analysis in the 10-K form. This results in MD&A's that are included by reference. These MD&A's are not used in this research.
- The company did not submit a 10-K form for the full period. This can be due to:
    - Entering or exiting the U.S. market
    - Going public or private
    - Mergers and acquisitions
    - Bankruptcy
    - Changing the company name
    - Change from or to the 10-K Small Business form, which was not successfully extracted and thus not present in the data.
    - The company submitted a 10-K Transition form, which was excluded from the large dataset

6.3 Coherence Score

Coherence, the independent variable, shows a large positive skew and 2294 (33%) of the MD&A's obtain a coherence score of zero. This was partly due to the fact that log(1) equals zero. This means that keywords occurring once are not integrated in the score. That is unwanted, because it eliminates valuable information and increases the chances of one of the elements obtaining a zero total score. If one coherence element obtains a zero score, the total coherence score is zero as well due to the multiplication of the element scores. These characteristics are decreased by taking the logarithm of twice the number of hits per keyword. Instead of log(keyword hits), log(2*keyword hits) is used. This has an effect because a keyword occurring once will obtain a value of log(2), roughly 0.3, instead of zero.

After this moderation 511 (7.3%) of the MD&A's still have a coherence score of zero. The large positive skew has somewhat decreased, but still exists. Therefore the approach in the analyses is to divide the

coherence score into equally sized low, medium and high coherence segments and then compare the high and low coherence groups to each other. The cut-off points for equal groups are situated at coherence scores of 0.01654 and 0.04284, indicative for the large skew.

The use of these coherence groups decreases the leverage and influence of outliers, because they are given the same 'value' as the other data points in their group. Therefore there is no need to remove outliers. To show more information about these groups, their descriptives are given in Table 16. As the groups in reality have assigned values of either zero or one, this gives little information about the data. Therefore the real data, which is not used in this way, is shown in Table 16.

Table 16 Descriptives of the coherence score data.

	N		Mean	Std. Deviation	Percentiles		
	Valid	Missing			25	50 (=Median)	75
Coherence Score Original	6923	0	0.0470	0.0690	0.0121	0.0271	0.0551
Low Coherence	2308	4615	0.0074	0.0054	0.0023	0.0077	0.0121
Medium Coherence	2307	4616	0.0280	0.0075	0.0216	0.0271	0.0342
High Coherence	2308	4615	0.1055	0.0940	0.0551	0.0750	0.1145

6.4 Premium Measures

In this section the dataset for the four premium variables EBIT margin, sales growth, return on assets and P/E ratio will be discussed and corrected.

6.4.1 Earnings Before Interest and Taxes margin

$$EBIT\ margin = \frac{ebit}{sales}$$

This metric is calculated by dividing EBIT by sales. This means that it is sensitive to total sales differences at the lower end of total sales. Many outliers appear in the region where total sales are limited. To improve reliability the dataset will be limited to companies with sales of over $10 million. 123 cases are removed. The first 30 cases have no sales and thus an EBIT margin of zero. The remaining 93 cases contain only 51 EBIT margins between -1 and 1. Twelve EBIT margins are smaller than -10, ranging to -409.

After removing these 123 cases, the maximum EBIT margin observed is .69, the lowest five cases range from -3.9 to -14.1. In order to limit the major influence of large absolute values, the values that are more than five standard deviations away from the mean, under -1.47 and above 1.54, are excluded. This amounts to nineteen cases, which all have negative values.

The original and cleaned data descriptives are given in Table 17.

Table 17 Descriptives of the EBIT margin data.

	N		Mean	Std. Deviation	Percentiles		
	Valid	Missing			25	50 (=Median)	75
EBIT margin original	6923	0	-11.79%	5.785	1.35%	4.86%	8.88%
EBIT margin cleaned	6781	142	4.60%	0.103	1.57%	4.98%	8.96%

6.4.2 Sales Growth

$$sales\ growth = \frac{sales\ current\ year}{sales\ previous\ year} - 1$$

This metric is calculated by dividing the current year sales by the previous year's sales. Therefore it is sensitive to previous year's sales differences when sales were low in the previous year. Inspecting the data we indeed find that many outliers appear in the region where previous year sales are limited. To improve reliability the dataset will be limited to companies with previous year sales of over $10 million. 221 cases are removed; 119 cases had either zero sales in the previous year or no previous year sales figure was available. 36 of the other 102 values were out of the -1 to 1.5 range.

Analysing the remaining extreme values, the five lowest values were found to be -1 and the five highest values are 12.96-22.59. It is apparent that the highest values have much more leverage than the lowest values. This amount of leverage is unwanted, but the data can still contain valuable and accurate information. With a cut-off at five standard deviations (5x.66) from the mean (.13), all cases with sales growth higher than 3.43 are deleted. Next to the large number of standard deviations from the mean it seems reasonable to treat companies with previous year sales of over $10 million that grow with more than 343% as outliers. 22 cases are removed in this process.

The original and cleaned data descriptives are given in Table 18.

Table 18. Descriptives of the sales growth data.

	N		Mean	Std.	Percentiles		
	Valid	Missing		Deviation	25	50 (=Median)	75
Sales growth original	6923	0	107.96%	65.009	-2.317%	5.57%	16.74%
Sales growth cleaned	6680	243	10.41%	0.318	-2.557%	5.75%	16.64%

6.4.3 Return on Assets

$$return\ on\ assets = \frac{net\ income}{total\ assets}$$

This metric is calculated by dividing the net income by the total assets. This means that it is very sensitive to small (absolute) differences in the denominator, total assets, when the denominator is small. Inspecting the data it is indeed found that many outliers appear in the region where total assets value is limited. To improve reliability the dataset will be limited to companies with total assets worth over $6 million. From that point there are significantly less extreme values, both in value and number. 66 cases were removed. In these 66 cases, 23 were outside the -1 to 1 range.

The remaining ROA values did not appear very extreme. The five highest ROA values are 1.65-2.44, the five lowest are -2.63 to -4.32 with a clear outlier of -14.74. The three cases with a ROA smaller than -4 are removed because their influence is very high; they are more than thirteen standard deviations separated from the mean.
Examining the items with an ROA of exactly zero, it was found that this was correlated with the current and previous year sales figures both being equal to zero in fifteen of the 21 cases. These fifteen cases were excluded, because the MD&A of a company that has not been in business in that year contains no (valuable) information.

The original and cleaned data descriptives are given in Table 19.

Table 19 Descriptives of the return on assets data.

	N		Mean	Std.	Percentiles		
	Valid	Missing		Deviation	25	50 (=Median)	75
ROA original	6923	0	-2.96%	0.9540	-2.22%	3.05%	7.37%
ROA cleaned	6840	83	-0.29%	0.2106	-2.08%	3.08%	7.39%

6.4.4 Price/Earnings Ratio

$$P/E\ Ratio = \frac{common\ shares\ outstanding * share\ price}{net\ income}$$

This metric is calculated by multiplying the number of common shares outstanding with the end of year share price and dividing the outcome by the net income. First of all, the dataset contains a large number of negative P/E ratios. This occurs when a company has a negative net income. Actually nothing conclusive can be said about the P/E ratio in these cases. Although exchanging the negative P/E ratio for a positive one might give us a reasonable order of magnitude estimate, the reliability of such a P/E ratio tremendously decreases. It is not justified to convert a loss to a profit here. Therefore all negative P/E ratios will be discarded first. The P/E ratio denominator is net income, which makes the P/E ratio sensitive to fluctuations around small net incomes. Inspecting the data it is indeed found that many outliers appear in the regions where the net income is small. To improve reliability the dataset will be limited to companies with net earnings above $0.35 million. In this process a total of 2,707 cases have been excluded, of which 2,220 have negative P/E ratios. This is a large number, but with the 4216 remaining cases the analyses can still be performed.

Now the top five highest P/E ratios are 606-11,319 (more than three standard deviations from the mean), the top five lowest are .0005-.016. Both can be called extreme. At the low end it is unreasonable to value a company at less than one year's profits and at the other end P/E ratios of in the thousands are unreasonable as well. These data points have a lot of leverage. Therefore the top five highest P/E ratios were removed, as well as all P/E ratios below one. Additionally, one case with current and previous year sales figures both being equal to zero was excluded, because the MD&A of a company that has not been in business in that year is not of interest to us. In this step 36 cases were removed, adding up to a total of 2,743 removed cases

The original and cleaned data descriptives are given in Table 20.

Table 20 Descriptives of the P/E ratio data.

	N		Mean	Std.	Percentiles		
	Valid	Missing		Deviation	25	50 (=Median)	75
P/E ratio original	6923	0	12.485	767.458	-0.398	10.705	19.462
P/E ratio cleaned	4180	2743	24.516	34.669	11.886	16.843	24.736

6.4.5 Correlations amongst Premium Measures

Now that all the premium measures have been described and cleaned it is important to look at their correlations. Very high correlations could indicate that one of the premium measures is obsolete. Looking at the correlations in

Table 21, there are two premium measures that are moderately correlated: ROA and EBIT margin. Interesting is the fact that P/E ratio is negatively correlated to ROA and EBIT margin. This could indicate a trade-off between P/E ratio and ROA/EBIT margin, but the correlations are low to negligible, which makes such an indication highly tentative. It is concluded that the current premium measures are fine. It could be the case that the ROA and EBIT margin measures show similar results. This could then partly be attributed to their moderate correlation.

Table 21 Correlations between the four premium measures. ROA and EBIT margin are moderately correlated. Suffix ** indicates that the correlation is significant at the 0.01 level (2-tailed).

		Sales growth	ROA	EBIT margin	P/E ratio
Sales growth	Pearson Correlation	1	.081**	.063**	.090**
	Sig. (2-tailed)		.000	.000	.000
	N	6680	6650	6648	4140
ROA	Pearson Correlation	.081**	1	.566**	-.235**
	Sig. (2-tailed)	.000		.000	.000
	N	6650	6840	6759	4174
EBIT margin	Pearson Correlation	.063**	.566**	1	-.116**
	Sig. (2-tailed)	.000	.000		.000
	N	6648	6759	6781	4175
P/E ratio	Pearson Correlation	.090**	-.235**	-.116**	1
	Sig. (2-tailed)	.000	.000	.000	
	N	4140	4174	4175	4180

6.5 Other Relevant Factors

The other factors that can influence premium and that can be included in the analyses are presented here. Size, diversification, industry and fiscal year will be covered.

6.5.1 Size

The size of a firm is measured by its sales. The sales descriptives for the large dataset is shown in Table 22. The size differences in the sample are large. Using size as a scale variable would mean that the large relative differences between company size at the high and low end are expected to be reflected in the size effect. However, it is expected that there are several size thresholds above or below which scale advantages or disadvantages take place. That is why the decision was made to work with three equally sized groups (small sales, medium sales and large sales) with cut-off points at $206.349 and $888.589 million.

Table 22 Descriptives of the size data.

	N		Mean	Std. Deviation	Percentiles		
	Valid	Missing			25	50 (=Median)	75
Sales (million $)	6923	0	2241.4	6436.5	129.1	437.8	1389.8

6.5.2 Diversification

The large sample contains companies that are active in one to six segments. 66 items were active in zero segments according to the data. These cases were excluded. The degree of diversification is shown in Figure 21. More than half of the sample is active in one segment only. It is assumed that the difference between two, three, four or six segments is not very significant for the coherence premium relationship. To keep the sub sample size large enough the sample was divided into two groups: undiversified (one segment) and diversified (more than one segment).

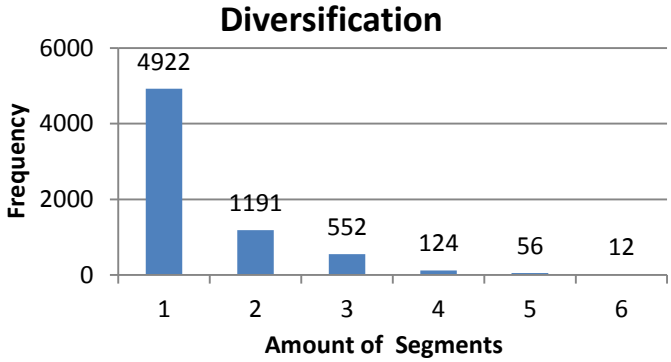

Figure 21 Histogram of the diversification of companies in the dataset, ranked by number of segments with activity.

6.5.3 Industry

From the literature it became apparent that the industry a company competes in has a significant influence on the premium. Therefore this is an important factor to take into account. There are several options to do this. It can be done with the FamaFrench-48 industry classification or with the SIC-code. Both will be discussed and the two-digit SIC-code is chosen, because it is more specific and is widely used.

FamaFrench-48 Industries

This is a classification into 48 industries, invented by Eugene Fama and Kenneth French. They constructed it from the SIC-codes. The FamaFrench classifications were constructed by Joost Impink from the SIC-codes the SEC uses. As the data consists of eight FamaFrench-48 industries, this is the initial logical choice. However, looking at Figure 22, it can be seen that two segments are very small, one is small and one is very large. This means that three segments could have difficulties achieving significant relations due to the small sample size and one segment risks the existence of very different firms within it.

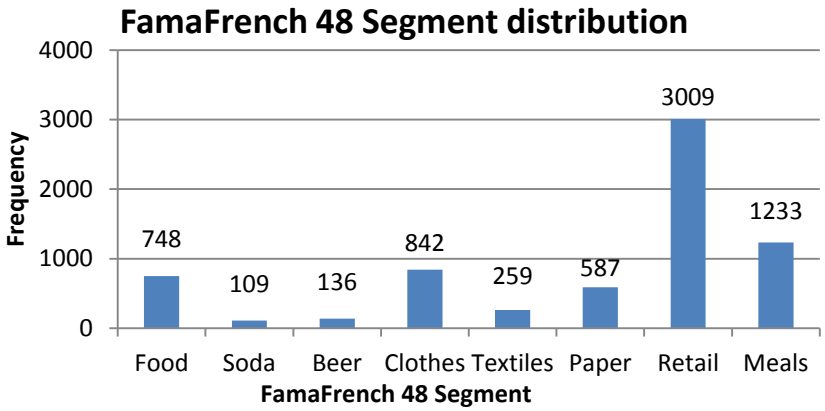

Figure 22 FamaFrench-48 industries distribution over the data.

SIC-Code (two-digit)

The SIC-code could provide an alternative. There are eighteen two-digit SIC codes present in the large sample. Their distribution is shown in Figure 23. The eighteen different SIC-codes make this approach more specific than the FamaFrench-48 industries classification. There are six SIC-codes that have a sample size that is expected to lead to significant results, if these are present. This is one more than for the FamaFrench-48 industries classification. Together with the added specificity and the fact that the

SIC-codes are very often used as the classification method in research, this leads to the choice to use two-digit SIC-codes as industry classification method.

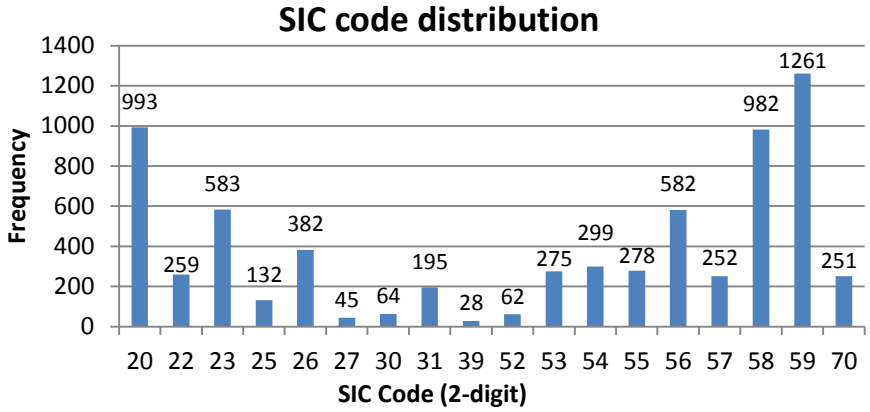

Figure 23 Two-digit SIC code distribution in the data.

6.5.4 Fiscal Year

Last there is the fiscal year, which distribution is shown in Figure 24. The distribution is fairly even, except for 2010, which has significantly less entries. The fiscal year has no direct influence on the relationship between coherence and premium. However, it could cause problems when two samples with a significant number of different fiscal year entries are compared and the different years are in different parts of the economic cycle. Looking at this data and knowing that most of the data, 79.36%, comes from companies that are present in the data for at least five years (see Figure 20), this effect is likely to be limited. Including the fiscal year as correction factor would decrease the sample size for specific combinations with size, industry and diversification. This would lead to a low chance of finding significant relations. Therefore the fiscal year is not taken into account from this point onwards.

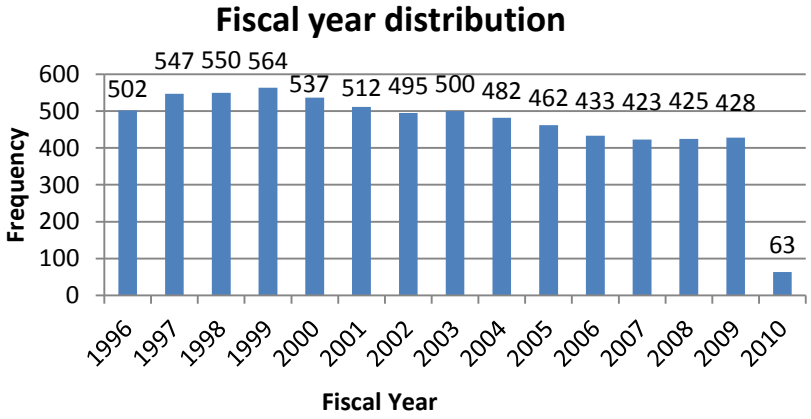

Figure 24 Fiscal Year distribution in the dataset.

6.6 Conclusion

In this chapter the data in general, coherence score data, premium measure data and other relevant factor data was inspected and cleaned cautiously, to prevent the deletion of valuable, accurate information. The coherence score data was highly skewed and non-normal; therefore it was decided to split it into three groups: low, medium and high coherence. The other relevant factor year is not taken into account in the analyses, because there is no rationale for its effect on the coherence premium relationship. Furthermore its distribution is very even.

In the next chapter, the high and low coherence groups will be compared by means of an ANOVA, which evaluates the statistical significance of differences between the means of the two groups.

7. Results

In this chapter the analyses described in the methodology section are performed with the large sample, which includes almost 7,000 data points for the consumer and retail industries. These large-scale analyses enable us to test the hypothesis; are coherence and premium positively correlated? First the one-way ANOVA is performed on the full large sample to see whether the high and low coherence group means differ significantly for the premium measures used. These results will be discussed in the first section. Then, in Section 0, the ANOVA is repeated for several subsets of the large sample to test for the effects of coherence in companies of different sizes, industries and diversification. Subsequently four case studies are performed, in which the combined effects of these tests are evaluated.

7.1 Analysis of Variances Full Sample

With the One-Way ANOVA analysis the means of the high and low coherence groups were compared for the premium measures ROA, P/E ratio, EBIT margin and sales growth. The full results are shown in Appendix V.A; a summary of the results is shown in Table 23.

- The means for all of these premium measures were higher in the high coherence group than in the low coherence group. The high coherence group thus outperforms the low coherence group on their means for every premium measure. Going from the low coherence to the high coherence group P/E ratio increases with 2.4, from 23.1 to 25.5, the ROA increases with 1.08%, from -.81% to .29%, EBIT margin increases with 0.57%, from 4.46% to 5.03% and sales growth increases with 2.6%, going from 9.5% to 12.1%.

- The ANOVA is significant at the 0.05 (and 0.01) level for sales growth and significant at the 0.1 level for the other premium measures. This means that the null hypothesis of equal means is rejected and it is concluded that sales growth and the other measures differ significantly between high and low coherence group at the 0.05 and 0.1 level respectively. The result for sales growth is very strong as it is significant up to the .007 level and the difference between the means is substantial.

- Inspecting Levene's Test of Homogeneity of Variances, it can be seen that for EBIT margin the high and low coherence groups do not have equal variances at a significance level of 0.05. This means that not all of the ANOVA assumptions are met and the means of EBIT margin are compared with the Kruskal-Wallis test, which is the non-parametric ANOVA. These results, stated in Appendix V.A, indicate that the means of EBIT margin differ at the .05 significance level.

Table 23 One-way ANOVA results for the full large sample. The suffixes *, ** and *** indicate that differences between the high and low coherence groups are significant at the .1, .05 and .01 significance level respectively. Suffix **** indicates that the variances between the two coherence groups are unequal at the .05 significance level.

FULL SAMPLE		Descriptives				ANOVA	Levene
		N	Mean	Std. Deviation	Std. Error	Sig.	Sig.
Sales growth	Low Coherence	2221	.0946	.3025	.00642	.007***	.210
	High Coherence	2221	.1209	.3408	.00723		
ROA	Low Coherence	2263	-.0081	.2281	.00479	.087*	.067
	High Coherence	2288	.0029	.2033	.00425		

EBIT margin	Low Coherence	2254	.0446	.0837	.00176	.069*	.000****
	High Coherence	2248	.0503	.1222	.00258		
P/E ratio	Low Coherence	1271	23.102	34.933	.9799	.077*	.332
	High Coherence	1554	25.458	35.478	.9000		

Discussion of Results

It is clear that sales growth is increased in the high coherence group. For the other measures there is an indication that this is the case, but the difference is less significant and smaller in (relative) magnitude. EBIT margin comes close to the 0.05 level of significance, but the absolute difference between the low and high coherence groups is limited. It must be noted that the variance inequality of the two coherence group means can influence the ANOVA results. For the P/E ratio measure it is important to note that, as mentioned in the premium part of Section 4.1.2, the P/E ratio is not a function of current performance only. This means that the interpretation of the P/E ratio difference in the means cannot be fully attributed to current year coherence. ROA has the lowest significance.

An explanation for the increased significance of the sales growth and EBIT margin metrics can be found in the fact that they are the fundamental value drivers. Profits are driven by sales and EBIT margin; future profits are driven by future sales and EBIT margin. Therefore sales growth and EBIT margin have a large impact on company valuations. Maybe the effect of coherence on the other premium measures is limited, because they are further away from this fundamental value creation. ROA only measures the return in relation to assets. It does not take into account labour, energy, materials etc. The P/E ratio is even further away because it is not a function of the current performance only.

Inspecting the absolute values obtained for the four premium values, the ROA and EBIT margin values appear to be low. According to Young and Basarab (2009), an EBIT margin of 5% is considered low, 10-12% is medium and 15+% is high. For ROA 3% is considered low, 5-7% is medium and 10+% is high. Using this classification, the EBIT margin values are low and the ROA values are very low. This could be caused by the effect of large negative values in the data. In this case a more strict protocol to remove outliers might improve this. Concerning sales growth 8-10% is nice (Young and Basarab 2009). The sales growth in the large sample is therefore regarded to be a little higher than normal.

The standard deviation from the mean for the high coherence groups is generally higher than for the low coherence groups. This can be explained by the large skew of the data. As most coherence scores are (much) closer to zero than to one, the low coherence group consists of coherence scores that lie close together. The large coherence group consists of coherence scores in a larger range, from 0.04284 to 1 to be specific.
As can be seen in Appendix V.A, the variance within group exceeds the variance between groups by far. This means that a large sample size is needed to obtain significant results. Fortunately the large sample is large enough to obtain several significant relationships. Although the results obtained are sound they are subject to potential biases. The effects of size, industry and diversification are for instance not taken into account; these will be investigated in the next section.

Taking into account the choices in the methodology section these results indicate that more emphasis placed on the coherence elements in the MD&A is positively related to performance, at least for the four measures that were used. As the emphasis placed on coherence in the MD&A is several steps away from measuring the actual coherence, it is quite remarkable that these results show up. In that sense, these results with a derivative of coherence really substantiate Leinwand and Mainardi (2010). On the other hand, the derivative can show a stronger effect than the original measure as well.

7.2 Analysis of Variances Corrected for Other Relevant Factors Individually

Now that it has been shown that coherence has an effect on the full large sample, its effect in several subsamples of the large sample will be assessed. In this section the effect of diversification, size and industry on the relationship between coherence and premium are evaluated. One-way ANOVA's will be performed for diversified, undiversified, small, medium and large firms as well as for firms within a two-digit SIC-code.

7.2.1 Diversification

An ANOVA was performed with two subsamples of the large sample: undiversified firms, active in one segment, and diversified firms, active in two to six segments. Beforehand it was hypothesized that diversification would moderate the relationship between coherence and premium. It is concluded that this is indeed the case. A summary of the one-way ANOVA results is shown in Table 24.

In undiversified firms the high coherence group has a significantly higher sales growth and P/E ratio. Sales growth, highly significant in the full sample, is even more significant for undiversified firms, but is insignificant for diversified firms. For diversified firms the high coherence group has a significantly higher ROA than the low coherence group. However, the variances are significantly not equal. Although not significant, it stands out that for sales growth and P/E ratio the low coherence group outperforms the high coherence group in the diversified firms group. This is the inverse of the significant relationships found for undiversified firms. As the sample size is smaller, it is not surprising to find less significant differences. For diversified firms, the ROA and EBIT margin differences and significance are increased, compared to the full sample.

Undiversified firms have more low coherence items, whereas diversified firms have more high coherence items. An explanation for this could be found in the survival rate. Diversified firms will generally be larger than undiversified firms, because it takes time to expand to new markets. Companies need time to grow large, but only the most effective companies, with high coherence, will survive long enough to reach that size. However, the correlation between size and diversification in this dataset is limited, around 0.15 at a significance level of 0.05.

Looking at the absolute values, the ROA and EBIT margin values are almost 2% higher for diversified firms. The P/E ratio is, on average, only 0.09 smaller for the undiversified firms, which constitutes a negligible diversification premium. Diversified firms have a very small 0.2% higher sales growth on average.

It is concluded that diversification makes a significant difference. Undiversified firms exhibit stronger relationships for sales growth and P/E ratio. Diversified firms exhibit stronger relationships for ROA and EBIT margin and show insignificant but inversed differences for sales growth and coherence.

Table 24 Summary of ANOVA diversification results. The suffixes *, ** and *** indicate that differences between the high and low coherence groups are significant at the .1, .05 and .01 significance level respectively. Suffix **** indicates that the variances between the two coherence groups are unequal at the .05 significance level.

DIVERSIFICATION			Descriptives				ANOVA	Levene
			N	Mean	Std. Deviation	Std. Error	Sig.	Sig.
Undiversified	**sales growth**	Low Coherence	1766	.0868	.2903	.00691	.000***	.004****
		High Coherence	1462	.1318	.3599	.00941		
	ROA	Low Coherence	1796	-.0085	.2226	.00525	.849	.685
		High Coherence	1505	-.0069	.2322	.00599		
	EBIT margin	Low Coherence	1793	.0423	.0860	.00203	.806	.000****

		High Coherence	1476	.0432	.1318	.00343		
	P/E ratio	Low Coherence	1015	22.455	31.933	1.002	.013**	.047****
		High Coherence	976	26.365	38.430	1.230		
Diversified	**sales growth**	Low Coherence	455	.1249	.3445	.01615	.182	.023****
		High Coherence	759	.0997	.2998	.01088		
	ROA	Low Coherence	467	-.00651	.2484	.01149	.008***	.001****
		High Coherence	783	.02187	.1290	.00461		
	EBIT margin	Low Coherence	461	.05356	.0733	.00342	.059*	.096
		High Coherence	772	.06370	.1002	.00361		
	P/E ratio	Low Coherence	256	25.67008	44.885	2.805	.509	.093
		High Coherence	578	23.92805	29.807	1.240		

7.2.2 Size

An ANOVA was performed with three subsamples of the large sample: small firms, with sales below $206.349 million, medium sized firms, with sales between $206.349 and $888.589 million, and large firms with sales of over $888.589 million. Initially, it was hypothesized that size would moderate the relationship between coherence and premium. It can be concluded that this effect is present in small firms, large firms and to a lesser extent to medium sized firms. A summary of the one-way ANOVA results is shown in Table 25.

In all but one of the cases where significant differences, at the 0.05 and 0.1 level, were found, the high coherence group outperforms the low coherence group. However, except for ROA in large companies and sales growth in medium sized companies, they all exhibit significantly unequal variances. This may affect the results.

- Within small firms the differences are significant and enlarged for sales growth and P/E ratio. The relationship of coherence with EBIT margin (significant) and ROA (insignificant) are reversed. This could be related to the start-up phase requiring large investments.
- The medium sized firms group only deviates from the full large sample findings in the insignificance of the ROA and EBIT margin relationships with coherence.
- Large companies show large differences between the low and high coherence group means for ROA and EBIT margin. Sales growth and P/E ratio differences became insignificant

It is insightful to note that the small company group mainly has low coherence items, the medium sized group has an approximately equal distribution and the large group mainly consists of high coherence items.

Looking at the absolute values, the ROA increases from -6.1% to 1.2% to 4.5% going from small to medium to large firms. EBIT margin is very low for the small size group (0.9%) compared to the medium (6.1%) and large (7.3%) groups. The differences for sales growth (medium > small > large) and P/E ratio (medium > large > small) are smaller. The direct effect of the size groups on the premium measure values, especially with ROA and EBIT margin, is large.
Generalised, large companies perform better than smaller companies. This could be caused by the fact that not all companies survive. As it takes time to reach a certain company size, larger companies will usually have lasted longer. The same argument can be applied to the fact that larger companies have a larger fraction of high coherence items. This coherence was needed to survive.

All in all, the differences between the coherence groups for the size groups are significant. Differences and significance, compared to the full large sample, are enlarged for sales growth and P/E ratio within smaller companies and for ROA and EBIT margin for larger companies. The reverse is true for the other

premium measures and the medium size category only deviates from the general results by lower significances. These results clearly differentiate between the effects of coherence in small and large companies. Therefore size is a factor for which coherence results should be corrected.

Table 25 Summary of one-way ANOVA size results. The suffixes *, ** and *** indicate that differences between the high and low coherence groups are significant at the .1, .05 and .01 significance level respectively. Suffix **** indicates that the variances between the two coherence groups are unequal at the .05 significance level.

SIZE			Descriptives				ANOVA	Levene
			N	Mean	Std. Deviation	Std. Error	Sig.	Sig.
Small	**Sales growth**	Low Coherence	835	.0819	.3610	.01249	.008***	.021****
		High Coherence	656	.1362	.4255	.01661		
	ROA	Low Coherence	856	-.0536	.3044	.01040	.278	.539
		High Coherence	709	-.0700	.2912	.01094		
	EBIT margin	Low Coherence	847	.0200	.1055	.00363	.001***	.000****
		High Coherence	669	-.0044	.1850	.00715		
	P/E ratio	Low Coherence	402	21.567	31.491	1.571	.038**	.036****
		High Coherence	342	27.227	42.564	2.302		
Medium	**Sales growth**	Low Coherence	763	.1032	.2708	.00980	.067*	.023****
		High Coherence	742	.1316	.3281	.01205		
	ROA	Low Coherence	780	.0094	.1742	.00624	.533	.115
		High Coherence	753	.0147	.1598	.00583		
	EBIT margin	Low Coherence	780	.0602	.0698	.00250	.583	.846
		High Coherence	753	.0622	.0771	.00281		
	P/E ratio	Low Coherence	436	22.926	28.565	1.368	.051*	.093
		High Coherence	517	27.440	40.373	1.776		
Large	**Sales growth**	Low Coherence	623	.1010	.2482	.00995	.884	.003****
		High Coherence	823	.0990	.2677	.00933		
	ROA	Low Coherence	627	.0324	.1372	.00548	.000***	.145
		High Coherence	826	.0548	.1056	.00367		
	EBIT margin	Low Coherence	627	.0584	.0526	.00210	.000***	.000****
		High Coherence	826	.0836	.0624	.00217		
	P/E ratio	Low Coherence	433	24.705	42.902	2.062	.442	.012****
		High Coherence	695	23.114	26.542	1.007		

7.2.3 Industry

Industry is operationalized as two-digit SIC code. An ANOVA was performed with eighteen two-digit SIC code subsamples of the large sample. It was hypothesized that industry would moderate the relationship between coherence and premium. It can be concluded that for some industries there are clear exceptions to the general trend in the full large sample. Industry thus does moderate the coherence premium relationship, at least for some industries. A summary of the one-way ANOVA results is shown in Table 26. The full ANOVA results can be found in Appendix X.

Excluded from this overview because at least one of the coherence groups consisted of less than fifteen items or the sum of the two groups consisted of less than 40 items:

- SIC code 25, furniture and fixtures;
- SIC code 27, printing, publishing and allied industries;
- SIC code 30, rubber and miscellaneous plastics;
- SIC code 39, miscellaneous manufacturing industries;
- SIC code 52, building materials, hardware, garden supply and mobile home dealers.

This was done to prevent conclusions to be drawn based on a very limited number of data points.

Excluded from this overview because no significant relationships were found:

- SIC-code 56, apparel and accessory stores;
- SIC-code 57, home furniture, furnishings and equipment stores.

In fifteen of the nineteen two-digit SIC code – premium measure combinations that have significant differences, the high coherence group outperforms the low coherence group.

- **SIC-code 20; food and kindred products**: The high coherence group outperforms the low coherence group for ROA and EBIT margin at the 0.05 and 0.10 significance level respectively.
- **SIC-code 22; textile mill products**: The high coherence group outperforms the low coherence group for ROA at the 0.10 significance level. The variances are unequal here.
- **SIC-code 23; apparel and other finished products made from fabrics and similar material**: The high coherence group outperforms the low coherence group for ROA and EBIT margin at the 0.05 significance level. The variances are significantly unequal for ROA.
- **SIC-code 26; paper and allied products**: The high coherence group does not outperform the low coherence group for sales growth at the 0.05 significance level. The difference is quite large, 4.5% for the high coherence group compared to 12.0% for the low coherence group. The variances are unequal here.
- **SIC-code 31; leather and leather products**: The high coherence group outperforms the low coherence group for EBIT margin at the 0.10 significance level, but with unequal variances.
- **SIC-code 53; general merchandise stores**: The high coherence group outperforms the low coherence group for EBIT margin and the low coherence group outperforms the high coherence group P/E ratio at the 0.01 significance level, both with unequal variances. The difference for the P/E ratio is very large, 18.3 for the high coherence group compared to 31.8 for the low coherence group.
- **SIC-code 54; food stores**: The high coherence group outperforms the low coherence group for sales growth at the 0.01 significance level. This difference is large, 1.0% for the low coherence group and 10.7% for the high coherence group.
- **SIC-code 55; automotive dealers and gasoline service stations**: The high coherence group outperforms the low coherence group for sales growth at the 0.01 significance level, with unequal variances. Their numbers and difference are huge, 15.6% for the low coherence group and 37.7% for the high coherence group.
- **SIC-code 58; eating and drinking places**: The high coherence group outperforms the low coherence group for sales growth, EBIT margin and P/E ratio at the 0.05 significance level, with unequal variances for EBIT margin only.
- **SIC-code 59; miscellaneous retail**: The high coherence group outperforms the low coherence group for sales growth, but the low coherence group outperforms the high coherence group for ROA and EBIT margin, all significant at the 0.01 significance level with unequal variances. Sales growth differs between 14.9% and 23.9% for the low and high coherence group respectively. ROA and EBIT margin are worse for the high coherence group: -8.7% and -1.7% versus -1.9% and 4.5% respectively.
- **SIC-code 70; hotels, rooming houses, camps and other lodging places**: The high coherence group outperforms the low coherence group for sales growth and EBIT margin at the 0.05 significance level with unequal variances for sales growth. Sales growth differs widely, between 5.7% and 32.8% for the low and high coherence group respectively

The differences between the SIC-codes high and low coherence groups resemble the differences found for the full sample. However, there are some distinct exceptions in some SIC-codes. SIC-code 59,

miscellaneous retail is the largest exception; two coherence premium relationships are reversed compared to the full large sample and significant at the 0.01 significance level. This could be due to the specificity of the industry, where coherence apparently does not lead to a premium or is overshadowed by other factors of influence. It could be that the samples contain unevenly distributed items of diversified, undiversified, small, medium and large firms. It could also be the case that there are other factors responsible for these deviations. For SIC-code 59 an extra explanation could be found in the 'miscellaneous' part, which indicates that it might be an inconsistent sample, which includes very different, retail related businesses that could not be classified elsewhere.
Looking at the absolute values of the premium measures, there is a lot of variability in their levels for the different SIC-codes.
All in all, the differences between the SIC-codes high and low coherence groups generally resemble the differences found for the full sample. However, there are some distinct exceptions in some SIC-codes. This makes it an important determinant of firm performance, which has to be taken into account.

Table 26 Summary of one-way ANOVA two-digit SIC-code industry results. The suffixes *, ** and *** indicate that differences between the high and low coherence groups are significant at the 0.10, 0.05 and 0.01 significance level respectively. Suffix **** indicates that the variances between the two coherence groups are unequal at the .05 significance level.

			Descriptives				ANOVA	Levene
Two -digit SIC CODE			**N**	**Mean**	**Std. Deviation**	**Std. Error**	**Sig.**	**Sig.**
20	**sales growth**	Low Coherence	252	0.0910	0.3066	0.0193	.674	.145
		High Coherence	338	0.0814	0.2458	0.0134		
	ROA	Low Coherence	258	-0.0001	0.2072	0.0129	.031**	.121
		High Coherence	348	0.0301	0.1362	0.0073		
	EBIT margin	Low Coherence	259	0.0467	0.0858	0.0053	.068*	.156
		High Coherence	341	0.0627	0.1198	0.0065		
	P/E ratio	Low Coherence	130	25.9476	33.7512	2.9602	.999	.485
		High Coherence	244	25.9452	29.7503	1.9046		
22	**sales growth**	Low Coherence	76	-0.0127	0.2046	0.0235	.547	.272
		High Coherence	103	0.0067	0.2183	0.0215		
	ROA	Low Coherence	78	-0.0668	0.3385	0.0383	.074*	.000****
		High Coherence	105	-0.0027	0.1192	0.0116		
	EBIT margin	Low Coherence	78	0.0460	0.0927	0.0105	.806	.724
		High Coherence	105	0.0411	0.1557	0.0152		
	P/E ratio	Low Coherence	33	13.4896	14.2069	2.4731	.103	.230
		High Coherence	58	21.1968	24.6180	3.2325		
23	**sales growth**	Low Coherence	173	0.0692	0.3630	0.0276	.400	.615
		High Coherence	198	0.0994	0.3272	0.0233		
	ROA	Low Coherence	174	-0.0393	0.3544	0.0269	.016**	.001****
		High Coherence	200	0.0281	0.1578	0.0112		
	EBIT margin	Low Coherence	175	0.0398	0.0909	0.0069	.000***	.354
		High Coherence	198	0.0760	0.0775	0.0055		
	P/E ratio	Low Coherence	93	23.8215	37.3637	3.8744	.138	.016****
		High Coherence	141	18.5082	16.3950	1.3807		
26	**sales growth**	Low Coherence	77	0.1202	0.2978	0.0339	.014**	.000****
		High Coherence	176	0.0445	0.1816	0.0137		
	ROA	Low Coherence	82	-0.0133	0.1362	0.0150	.149	.286
		High Coherence	181	0.0143	0.1460	0.0109		
	EBIT margin	Low Coherence	82	0.0573	0.0824	0.0091	.220	.725
		High Coherence	179	0.0731	0.1022	0.0076		
	P/E ratio	Low Coherence	37	21.7761	23.6614	3.8899	.685	.160
		High Coherence	127	20.0119	23.1467	2.0539		
31	**sales growth**	Low Coherence	50	0.0867	0.5549	0.0785	.813	.007****
		High Coherence	97	0.1015	0.1998	0.0203		
	ROA	Low Coherence	48	0.0437	0.1816	0.0262	.168	.009****
		High Coherence	97	0.0793	0.1244	0.0126		

	EBIT margin	Low Coherence	49	0.0709	0.1218	0.0174	.096*	.013****
		High Coherence	97	0.0993	0.0812	0.0082		
	P/E ratio	Low Coherence	31	15.8957	9.8526	1.7696	.606	.341
		High Coherence	83	17.0749	11.1655	1.2256		
53	**sales growth**	Low Coherence	64	0.0768	0.2790	0.0349	.889	.007****
		High Coherence	107	0.0722	0.1461	0.0141		
	ROA	Low Coherence	64	-0.0152	0.1463	0.0183	.269	.448
		High Coherence	107	0.0208	0.2333	0.0226		
	EBIT margin	Low Coherence	64	0.0203	0.0732	0.0091	.002***	.042****
		High Coherence	107	0.0479	0.0410	0.0040		
	P/E ratio	Low Coherence	43	31.7549	39.1482	5.9700	.003***	.000****
		High Coherence	89	18.2764	10.5754	1.1210		
54	**sales growth**	Low Coherence	101	0.0098	0.1908	0.0190	.002***	.220
		High Coherence	95	0.1069	0.2387	0.0245		
	ROA	Low Coherence	102	-0.0159	0.1335	0.0132	.625	.724
		High Coherence	97	-0.0040	0.2028	0.0206		
	EBIT margin	Low Coherence	101	0.0208	0.0268	0.0027	.372	.170
		High Coherence	96	0.0306	0.1076	0.0110		
	P/E ratio	Low Coherence	38	20.7124	15.5841	2.5281	.157	.062
		High Coherence	66	34.2338	57.1985	7.0406		
55	**sales growth**	Low Coherence	133	0.1564	0.3241	0.0281	.004***	.001****
		High Coherence	41	0.3771	0.6623	0.1034		
	ROA	Low Coherence	134	0.0022	0.1152	0.0100	.914	.486
		High Coherence	43	0.0045	0.1393	0.0212		
	EBIT margin	Low Coherence	133	0.0313	0.0432	0.0037	.124	.019****
		High Coherence	42	0.0444	0.0600	0.0093		
	P/E ratio	Low Coherence	90	21.9198	56.6021	5.9664	.879	.829
		High Coherence	32	23.5083	26.3562	4.6592		
58	**sales growth**	Low Coherence	415	0.0848	0.2429	0.0119	.035**	.340
		High Coherence	213	0.1300	0.2724	0.0187		
	ROA	Low Coherence	418	0.0018	0.1782	0.0087	.816	.336
		High Coherence	220	-0.0016	0.1688	0.0114		
	EBIT margin	Low Coherence	423	0.0429	0.0668	0.0033	.012**	.000****
		High Coherence	216	0.0594	0.0980	0.0067		
	P/E ratio	Low Coherence	220	20.7685	28.2569	1.9051	.011**	.084
		High Coherence	148	29.0339	33.4976	2.7535		
59	**sales growth**	Low Coherence	401	0.1494	0.3076	0.0154	.003***	.000****
		High Coherence	373	0.2388	0.5091	0.0264		
	ROA	Low Coherence	409	-0.0186	0.2864	0.0142	.002***	.001****
		High Coherence	403	-0.0865	0.3408	0.0170		
	EBIT margin	Low Coherence	400	0.0450	0.1010	0.0051	.000***	.000****
		High Coherence	381	-0.0169	0.1862	0.0095		
	P/E ratio	Low Coherence	246	27.8170	48.0075	3.0608	.202	.381
		High Coherence	217	33.0387	38.5586	2.6175		
70	**sales growth**	Low Coherence	96	0.0573	0.5027	0.0513	.005***	.003****
		High Coherence	64	0.3284	0.7090	0.0886		
	ROA	Low Coherence	106	-0.0226	0.2212	0.0215	.437	.062
		High Coherence	71	-0.0003	0.1191	0.0141		
	EBIT margin	Low Coherence	100	0.0818	0.1421	0.0142	.017**	.648
		High Coherence	69	0.1361	0.1452	0.0175		
	P/E ratio	Low Coherence	45	23.5700	23.6414	3.5243	.170	.120
		High Coherence	52	40.5688	79.4794	11.0218		

7.3 Analysis of Variances Corrected for Relevant Factors Together

Now that the individual influences of size, diversification and industry were shown, it is interesting to look at the effect of the combinations between them for the coherence premium relationship. To divide the full large sample into specific industry-size-diversification subsamples with a large enough size to do statistical tests with, only the three largest SIC-codes will be used. These are:

- SIC code 20, food and kindred products, 993 items;
- SIC code 58, eating and drinking places, 982 items;
- SIC code 59, miscellaneous retail, 1261 items;

An ANOVA was performed for which the results are summarised in Table 27. The full ANOVA results can be found in Appendix V.C.

Excluded from this overview because at least one of the coherence groups consisted of less than 10 items or the sum of the two groups consisted of less than 30 items:

- SIC 20 – medium – diversified
- SIC 58 – small – diversified – P/E ratio
- SIC 58 – medium – diversified
- SIC 58 – large – diversified
- SIC 59 – small – diversified – P/E ratio

This was done to prevent conclusions to be drawn based on a very limited amount of data. The limits are lowered compared to the previous section, because the larger number of variables taken into account results in smaller subsamples. The deletion of more data would have led to a loss off the representational value of this overview.

Excluded from this overview because no significant relationships were found:

- SIC 20 – small
- SIC 58 – medium – undiversified

In SIC 20, food and kindred products:

- **Medium – Undiversified**: The high coherence group outperforms the low coherence group significantly for ROA at the 0.1 level, but the low coherence group outperforms the high coherence group for EBIT margin at the 0.05 level with unequal variances between the groups.
- **Large – Undiversified**: The high coherence group outperforms the low coherence group significantly for ROA and EBIT margin at the 0.05 level, with unequal variances between the groups for EBIT margin.
- **Large – Diversified**: The high coherence group outperforms the low coherence group significantly for ROA and EBIT margin at the 0.05 level, with unequal variances between the groups for ROA.

In SIC 58, eating and drinking places:

- **Small – Undiversified**: The high coherence group outperforms the low coherence group significantly for sales growth and P/E ratio at the 0.05 level with unequal variances between the groups for both. Differences between the means are quite large; 10.0% for sales growth and 14.8 for P/E ratio.
- **Small – Diversified**: The low coherence group outperforms the high coherence group significantly for ROA and EBIT margin at the 0.05 level with unequal variances between the

groups for both. Differences between the means are large: 13.9% and 7.5% for ROA and EBIT margin respectively.

- **Large – Undiversified**: The high coherence group outperforms the low coherence group significantly for ROA and EBIT margin at the 0.05 level with unequal variances between the groups for both.

In SIC 59, miscellaneous retail:

- **Small – Undiversified**: The high coherence group outperforms the low coherence group significantly for sales growth but the low coherence group outperforms the high coherence group significantly for ROA and EBIT margin, all at the 0.05 level with unequal variances. Differences are large: 16.1% for sales growth, 16.9% for ROA and 10.8% for EBIT margin, where the latter two include mainly negative values.
- **Small – Diversified**: The high coherence group outperforms the low coherence group significantly for sales growth and ROA at the 0.10 and 0.05 level respectively, with unequal variances between the groups for ROA. Again large differences occur: 22% for sales growth and 28.5% for ROA, which is negative for both coherence groups.
- **Medium – Undiversified**: The high coherence group outperforms the low coherence group significantly for P/E ratio, but the low coherence group outperforms the high coherence group significantly for ROA and EBIT margin, all at the 0.05 level with unequal variances for P/E ratio. P/E ratio difference is 28.1.
- **Medium – Diversified**: The low coherence group outperforms the high coherence group significantly for EBIT margin at the 0.05 level.
- **Large – Undiversified**: The low coherence group outperforms the high coherence group significantly for EBIT margin at the 0.10 level and there is a small difference of 1.4% between the coherence groups.
- **Large – Diversified**: The low coherence group outperforms the high coherence group significantly for EBIT margin at the 0.05 level.

It is not easy to draw conclusions from these case studies. The results could be caused mainly by the specific three SIC-codes that were used as test cases. SIC 59, miscellaneous retail, consists of retail companies that could not be classified elsewhere. It could be the case that these are very different companies put together in SIC 59. What can be seen is that the data and differences between the coherence groups vary wildly. While some trends were distinguished in the single factor correction analyses, no clear trends were found in the data here. It could be that the combination of different individual factors caused this. The conclusion is that analysis at this detailed level needs more attention and possibly another approach.

Table 27 Summary of one-way ANOVA relevant factors together case studies. The suffixes *, ** and *** indicate that differences between the high and low coherence groups are significant at the .1, .05 and .01 significance level respectively. Suffix **** indicates that the variances between the two coherence groups are unequal at the .05 significance level.

Two-digit SIC-code					Descriptives				ANOVA	Levene
	Size	Diversification			N	Mean	Std. Deviation	Std. Error	Sig.	Sig.
20	Medium	Undiversified	sales growth	Low Coherence	50	0.0593	0.1668	0.0236	.168	.752
				High Coherence	68	0.1163	0.2530	0.0307		
			ROA	Low Coherence	52	0.0174	0.1300	0.0180	.091*	.056
				High Coherence	69	0.0460	0.0446	0.0054		
			EBIT margin	Low Coherence	52	0.0867	0.0801	0.0111	.011**	.000****
				High Coherence	69	0.0586	0.0369	0.0044		

			P/E ratio	Low Coherence	29	25.6288	23.6518	4.3920	.447	.458
				High Coherence	55	31.1759	35.1095	4.7342		
	Large	Undiversified	sales growth	Low Coherence	40	0.1141	0.1865	0.0295	.182	.291
				High Coherence	58	0.0643	0.1763	0.0232		
			ROA	Low Coherence	41	0.0366	0.0712	0.0111	.005***	.982
				High Coherence	59	0.0767	0.0668	0.0087		
			EBIT margin	Low Coherence	41	0.0643	0.0519	0.0081	.000***	.010****
				High Coherence	59	0.1399	0.0746	0.0097		
			P/E ratio	Low Coherence	28	20.5238	15.4931	2.9279	.377	.526
				High Coherence	52	24.6581	21.7939	3.0223		
		Diversified	sales growth	Low Coherence	40	0.1716	0.4453	0.0704	.136	.047****
				High Coherence	48	0.0655	0.1830	0.0264		
			ROA	Low Coherence	41	0.0300	0.0335	0.0052	.000***	.000****
				High Coherence	48	0.0763	0.0588	0.0085		
			EBIT margin	Low Coherence	41	0.0462	0.0385	0.0060	.000***	.196
				High Coherence	48	0.1011	0.0439	0.0063		
			P/E ratio	Low Coherence	17	35.2442	72.0222	17.4679	.184	.007****
				High Coherence	43	19.8066	15.4036	2.3490		
58	Small	Undiversified	sales growth	Low Coherence	177	0.0933	0.2755	0.0207	.033**	.028****
				High Coherence	63	0.1929	0.4141	0.0522		
			ROA	Low Coherence	177	-0.0208	0.1811	0.0136	.135	.203
				High Coherence	64	-0.0602	0.1778	0.0222		
			EBIT margin	Low Coherence	182	0.0181	0.0770	0.0057	.458	.123
				High Coherence	63	0.0094	0.0882	0.0111		
			P/E ratio	Low Coherence	86	18.5920	15.5714	1.6791	.004***	.006****
				High Coherence	33	33.3577	38.9439	6.7793		
		Diversified	sales growth	Low Coherence	38	0.0528	0.2359	0.0383	.971	.658
				High Coherence	12	0.0498	0.2904	0.0838		
			ROA	Low Coherence	38	0.0540	0.1059	0.0172	.002***	.001****
				High Coherence	17	-0.0849	0.2040	0.0495		
			EBIT margin	Low Coherence	38	0.0431	0.0210	0.0034	.002***	.000****
				High Coherence	14	-0.0317	0.1428	0.0382		
	Large	Undiversified	sales growth	Low Coherence	42	0.0370	0.1564	0.0241	.226	.122
				High Coherence	36	0.0732	0.0914	0.0152		
			ROA	Low Coherence	42	0.0309	0.1790	0.0276	.068*	.005****
				High Coherence	36	0.0871	0.0359	0.0060		
			EBIT margin	Low Coherence	42	0.0723	0.0517	0.0080	.000	.004****
				High Coherence	36	0.1500	0.0704	0.0117		
			P/E ratio	Low Coherence	29	17.1509	8.8396	1.6415	.135	.871
				High Coherence	35	20.9292	10.7531	1.8176		
59	Small	Undiversified	sales growth	Low Coherence	120	0.1628	0.3914	0.0357	.031**	.000****
				High Coherence	112	0.3236	0.7052	0.0666		
			ROA	Low Coherence	124	-0.0450	0.3345	0.0300	.001***	.003****
				High Coherence	132	-0.2137	0.4644	0.0404		
			EBIT margin	Low Coherence	120	0.0056	0.1484	0.0135	.000***	.000****
				High Coherence	113	-0.1138	0.2894	0.0272		
			P/E ratio	Low Coherence	66	37.3610	62.2327	7.6603	.302	.058
				High Coherence	42	27.0317	21.2303	3.2759		
		Diversified	sales growth	Low Coherence	22	0.0537	0.3871	0.0825	.053*	.828
				High Coherence	31	0.2744	0.4090	0.0735		
			ROA	Low Coherence	23	-0.3766	0.7408	0.1545	.033**	.001****
				High Coherence	37	-0.0923	0.2329	0.0383		
			EBIT margin	Low Coherence	18	-0.0097	0.1344	0.0317	.634	.462
				High Coherence	34	-0.0285	0.1344	0.0230		
	Medium	Undiversified	sales growth	Low Coherence	98	0.1714	0.2722	0.0275	.370	.040****
				High Coherence	95	0.2183	0.4373	0.0449		
			ROA	Low Coherence	99	0.0170	0.1590	0.0160	.048**	.096
				High Coherence	98	-0.0466	0.2748	0.0278		

		EBIT margin	Low Coherence	99	0.0619	0.0719	0.0072	.000***	.305
			High Coherence	98	0.0084	0.1030	0.0104		
		P/E ratio	Low Coherence	50	19.6169	12.1611	1.7198	.003***	.000****
			High Coherence	53	47.7387	63.8589	8.7717		
	Diversified	sales growth	Low Coherence	38	0.1535	0.1410	0.0229	.726	.094
			High Coherence	30	0.1730	0.3032	0.0554		
		ROA	Low Coherence	39	0.0067	0.1740	0.0279	.630	.632
			High Coherence	31	-0.0116	0.1334	0.0240		
		EBIT margin	Low Coherence	39	0.0873	0.0498	0.0080	.008***	.081
			High Coherence	31	0.0482	0.0706	0.0127		
		P/E ratio	Low Coherence	25	17.7666	7.2000	1.4400	.214	.021****
			High Coherence	18	26.8202	35.0430	8.2597		
Large	Undiversified	sales growth	Low Coherence	90	0.1213	0.2407	0.0254	.197	.102
			High Coherence	59	0.1783	0.2929	0.0381		
		ROA	Low Coherence	91	0.0403	0.0645	0.0068	.261	.019****
			High Coherence	59	0.0048	0.2898	0.0377		
		EBIT margin	Low Coherence	91	0.0598	0.0305	0.0032	.068*	.325
			High Coherence	59	0.0455	0.0638	0.0083		
		P/E ratio	Low Coherence	75	28.4110	60.8769	7.0295	.785	.323
			High Coherence	46	25.9007	15.4080	2.2718		
	Diversified	sales growth	Low Coherence	33	0.1714	0.3072	0.0535	.994	.779
			High Coherence	46	0.1707	0.4509	0.0665		
		ROA	Low Coherence	33	0.0308	0.0433	0.0075	.972	.823
			High Coherence	46	0.0304	0.0632	0.0093		
		EBIT margin	Low Coherence	33	0.0768	0.0462	0.0080	.014**	.218
			High Coherence	46	0.0523	0.0399	0.0059		
		P/E ratio	Low Coherence	25	29.7000	26.8415	5.3683	.918	.754
			High Coherence	42	29.0634	22.6264	3.4913		

7.4 Conclusion

With the developed coherence approximation evidence was found for a positive relationship between company coherence and premium. Sales growth is significantly positively related to coherence, the other premium metrics are positively related to coherence as well, but to a limited extent. For ROA, P/E ratio, EBIT margin and sales growth the high coherence group outperformed the low coherence group based on their means. The means and their significance for the four premium measures are shown in Table 28. Sales growth is substantially positively related to coherence; the other measures show limited positive relationships.

Table 28 One-way ANOVA results for the full large sample. The suffixes * and ** indicate that differences between the high and low coherence groups are significant at the 0.10 and 0.01 significance level respectively.

Premium measure	Coherence Group	Mean	Sig.
Sales growth	Low Coherence	9.46%	.007**
	High Coherence	12.09%	
Return on assets	Low Coherence	-0.81%	.087*
	High Coherence	0.29%	
Earnings before interest and taxes margin	Low Coherence	4.46%	.069*
	High Coherence	5.03%	
Price/Earnings ratio	Low Coherence	23.1	.077*
	High Coherence	25.5	

Subsequently the coherence effect was corrected for size, diversification and industry. This was done by testing the premium differences between groups of different industries, between groups of different

sizes (by sales) and between diversified and non-diversified firms. It was hypothesized that these relevant factors moderate the relationship between coherence and premium.

- Diversification makes a significant difference. Undiversified firms exhibit stronger relationships for sales growth and P/E ratio. Diversified firms exhibit stronger relationships for ROA and EBIT margin and show insignificant but inversed differences for sales growth and coherence.
- The size groups affect the differences between the coherence groups too. Differences and significance, compared to the full large sample, are enlarged for sales growth and P/E ratio within smaller companies and for ROA and EBIT margin for larger companies. The reverse is true for the other premium measures and the medium size category only deviates from the general results by lower significances. These results clearly differentiate between the effects of coherence in small and large companies.
- Industry (by two-digit SIC-codes) high and low coherence group differences generally resemble the differences found for the full sample. However, there are some distinct exceptions in several SIC-codes. This makes it an important determinant of firm performance, which has to be taken into account.

In general, large and diversified companies perform better than smaller and undiversified companies. The datasets of the large and/or diversified groups consist of a relatively high fraction of highly coherent items. The better performance and the dataset composition could be caused by the survival bias; only the coherent and successful companies grow to become large multinationals, active in many segments.
A related finding is that diversified and larger firm groups mainly consist of high coherence companies, whereas the undiversified and smaller firm groups have more low coherence firms. An explanation can be found in the survival rate. Diversified firms are generally larger than undiversified firms, because it takes time to expand to new markets. Companies need time to grow large, but only the most effective companies, which are (highly) coherent, will survive long enough to reach that size. This implies that coherence and firm survival are positively correlated.

8. Conclusions and Recommendations

In this chapter the research questions will be answered and discussed. The first section will conclude the research project by answering the research questions. In Section 8.2 reflection on the obtained results will be discussed. Subsequently the limitations of the results are explained in Section 8.3. The final section adds recommendations for further research, based on these limitations.

8.1 Conclusions

The aim of this study was to quantitatively test whether coherence has a positive relationship with premium. An additional goal was to analyse the potential of textual analysis. These goals were fulfilled successfully.

Coherence was defined as alignment between three elements in a firm: its core capabilities system, way to play and product and service portfolio. A positive relationship between coherence and premium was hypothesized from indications by Leinwand and Mainardi (2010), the literature review and strategy consulting practice.
Coherence was operationalized by multiplying the core capabilities system score, way to play score and product and service portfolio score with each other. Using textual analysis to count a set of predefined words, the emphasis on the three coherence elements in the MD&A of the SEC 10-K filing was measured, relative to the MD&A length and a peer group of consumer and retail companies. The count of coherence element related keywords in these MD&A's was taken as an approximation for the emphasis of the company on coherence and was subsequently related to premium measures. Premium was defined as above normal financial performance compared to peers. ROA, EBIT margin, sales growth and P/E ratio were used as financial performance metrics.

The main research question, '*What is the relationship between 'coherence' in a company and the 'premium' it delivers?',* can now be answered. Evidence was found that coherence in a company is positively related with the premium it delivers. However, except for sales growth the differences are small and the significance level is limited. For ROA, P/E ratio, EBIT margin and sales growth the high coherence group outperformed the low coherence group based on their means. Furthermore it was shown that for all metrics the means[19] differ between high and low coherence groups with a significance level of 0.10. For sales growth the differences are significant up to the 0.01 level. Sales growth is substantially positively related to coherence; the other measures show limited positive relationships.
An explanation for the increased significance of the sales growth and EBIT margin differences could be that they are the fundamental value drivers. The effect of coherence on the other premium measures could be limited, because they are further away from this fundamental value creation. ROA only measures the return in relation to assets. The P/E ratio is even further away because it is not a function of the current performance only.

Subsequently the coherence premium relationship was corrected for size, diversification or industry. This was done by testing the premium differences between groups of different industries, between groups of different sizes (by sales) and between diversified and non-diversified firms. It was hypothesized that these relevant factors moderate the relationship between coherence and premium.

[19] The premium measure means for the low and high coherence groups and their significance level are shown in Table 28 **One-way ANOVA results for the full large sample. The suffixes * and ** indicate that differences between the high and low coherence groups are significant at the 0.10 and 0.01 significance level respectively.**Table 28 on page 80.

- Diversification makes a significant difference. Undiversified firms exhibit stronger relationships for sales growth and P/E ratio. Diversified firms exhibit stronger relationships for ROA and EBIT margin and show insignificant but inversed differences for sales growth and coherence.
- The size groups affect the differences between the coherence groups too. Differences and significance, compared to the full large sample, are enlarged for sales growth and P/E ratio within smaller companies and for ROA and EBIT margin for larger companies. The reverse is true for the other premium measures and the medium size category only deviates from the general results by lower significances. These results clearly differentiate between the effects of coherence in small and large companies.
- Industry (by two-digit SIC-codes) high and low coherence group differences generally resemble the differences found for the full sample. However, there are some distinct exceptions in several SIC-codes. This makes it an important determinant of firm performance, which has to be taken into account.

8.2 Reflection on the Results

Although some evidence was found, the results are not very strong, especially when considering the large dataset of almost 7000 items that was used. This is attributed to the fact that coherence cannot be measured directly and the coherence measure developed is only a limited approximation of coherence. It does not (fully) measure all the aspects of coherence. Although the opposite could be true, looking at the literature, case studies and coherence profiler survey results, the current results are likely to become stronger, when a more accurate approximation is used.
Therefore the current results still substantiate that the coherence theory has the potential to explain part of the variance in company performance. In our literature research it was shown that the coherence theory incorporates many aspects from the different strategy schools. Therefore, it would be of additional value as an overarching/combinatory/ bridging theory.

Indirectly the results substantiate the research findings that were mapped onto the coherence premium framework. However, the literature mapped onto the framework consisted mostly of adjacent research or fractions of publications; it was often either very general or specific. Therefore the results cannot conclude anything tangible about the literature findings.

The limited 'mappability' says something about the state of the strategy research field. It mainly aims at finding aspects that affect firm performance. As long as the majority of the variance cannot be explained there will be less attention paid to the question why a certain aspect causes differences in performance, as this question usually comes after the discovery. In that sense, Leinwand and Mainardi (2012) provide decent reasoning behind the theory. However, this could become more convincing if it is expanded. Still, for the coherence theory to become widely accepted, stronger results will have to be obtained, preferably with multiple methods. Especially since the textual analysis method is unconventional, additional backup with a large well-designed (traditional) survey will have a large additional value.

8.3 Research Limitations

The (automated) measurement of actual coherence on a large scale was not feasible; therefore an approximation was developed using textual analysis. Results with actual coherence may differ from current findings. It is not known in what direction these results will deviate. However, looking at the literature, case studies, practical experience and coherence profiler survey results, the current results are likely to become stronger when a more accurate approximation is used. This study is subject to many limitations, which implies that the differences found between the means of low and high

coherence groups should be regarded with caution. In this section the limitations related to scope, methodology, data and execution will be covered.

The study was limited to quantitative textual analysis ofMD&A's from the 10-K SEC-filings of 9.974 consumer and retail companies. Apart from limitations imposed by the method itself, these effects will be limited. No significant bias is introduced by the sample selection as long as results are not one on one translated to other sectors.
An import point to make is that the causality between coherence and premium was not evaluated in this study. It was possible to measure the relationship between the two, but regarding the direction of the causal relationship no conclusions can be drawn.
'The Coherence Premium' by Leinwand and Mainardi (2010) was taken as a starting point which had a large influence on the research. This effect was strengthened, because the literature search offered limited additions and extensions to the definition of coherence. The research was thus confined to the relationship between coherence and premium, antecedents of coherence or pre-coherence factors/stages were not investigated. The developed coherence premium framework was developed without testing its validity a priori or a posteriori. The coherence elements of Leinwand and Mainardi (2010) were left intact and four financial measures were used for premium. The framework posed limits, because factors not (explicitly) present in the model were excluded and not measured. On the other hand, the number of factors absent in the broad model is limited. The four premium measures are well-established and reliable, but the addition of other, quantitative or qualitative, measures could lead to new insights or other results. With four measures the validity of the premium results obtained can be ensured.

An additional limitation was introduced by the test set used to help deciding on the methodology, which consisted of companies from 'The Coherence Premium'. This test set has a very large influence on the results. It is questionable, whether this set is representative of the full sample as it was limited to 37 data points. The decisions in the methodology section were based on it. The most important assumption made was that the emphasis on coherence elements shown in an MD&A is a decent proxy for coherence itself. Stating the right things in a document does not mean management will act on it. Even if that is the case, the entire company will have to pick up on the message as well. It is difficult to assess the validity of these aspects and thus this is an important limitation. The 10-K writers within the company have a large impact on the results and this could have a large effect on the results obtained.

In the methodology section the measurement of coherence and premium was developed. For premium the procedures were quite straightforward, but the operationalization of coherence is quite complicated. Using the test set to help in the decision making introduced a bias and it cannot be ensured that the optimal solution (based on the test set) was found as not every possible combination of options was tried. As the operationalization of coherence was quite difficult and complicated, this is a real limitation. It has the potential to have a large effect on the results obtained. Therefore the results obtained must be used with caution.

In the steps to operationalize coherence, a number of choices were made. They affect the results of our analyses. The following choices were made:

- Base the keyword longlist on literature, synonyms and personal additions
- Use only single keywords in the keyword longlist
- Test the keyword longlist with five CEO-letters
- Perform the coding procedure by noting 0-1 for irrelevant – relevant and assigning a rating of 1-3 for the strength

With help from the test set the following choices were made for the keyword shortlist:

- Include general keywords only
- Use a manual aggregation of like keywords
- Only use keywords that appear at least once in the test data
- Cut-off at 70% relevance and an average score (relevant items only) of two

The test set helped making decisions for the measurement as well:

- Multiplying the coherence element's scores
- Counting the log(2*keyword hits) per keyword
- Comparing the coherence score and premium measure in the same year

The dataset used is limited in size and available information. Therefore it was not possible to take into account more aspects than diversification, size, industry (and fiscal year) as additional factors that influence the coherence premium relationship. Although this is limiting, the dataset has a considerable size, which enabled us to do many analyses. Enlarging it will have little effect on the results. In the data cleaning process choices were made regarding outliers. Alternative boundaries for outliers could lead to different end results. The data was limited to companies with sales over $10 million, total assets over $6 million or net income above $0.35 million for the different premium measures. This procedure could have introduced a small bias towards larger companies. In order to look at companies beneath these threshold values other data should be used. It was unfortunate that the final data was not suitable for more sophisticated analyses than the ANOVA.

Concerning the execution it should be noted that the vast majority of the work done was performed by one person. This could introduce a bias and has put depth and width constraints on some of the process steps. The latter two constraints reduce the reliability of the steps taken as less data was used to reach to conclusions. This might have a considerable effect on the results.

In the next section many of these limitations will be converted into recommendations for future research.

8.4 Recommendations

The current research presented is a first step towards the quantitative testing of the coherence premium relationship. The results obtained are promising, but need refinement and additional strength. Recommendations will be given on the following topics: scope and validation, textual analysis, methodology, data, validation and future research opportunities.

8.4.1 Scope and Validation

Considering the scope limitations there are alternatives to quantitative textual analysis, which could also provide answers to the research questions posed. Combining the results of quantitative textual analysis with other (qualitative) research methods like surveys and case studies can add additional strength to the conclusions presented here. This method of triangulation (Jick, 1979) derives its strength from the different types of research that all point in the same direction. This decreases the need for very strong individual research results.

- First of all, it is recommended to do more research using textual analysis to obtain stronger and more reliable results.
- Second, a scientifically set up coherence survey could provide additional validation, because its methodology and biases are inherently different from textual analysis.

- Additionally, this could be used (potentially together with textual analysis) as a tool to assess the coherence of prospective clients at Booz & Company.

The scope could also be extended towards pre-coherence, determinants of coherence elements, or after-coherence elements. The framework used here can be replaced by another one. It would be very beneficial to test whether the assumptions behind the model hold. If they do, this will help validating the coherence premium theory.
Comparing data from different points in time or another alternative approach could allow for conclusions regarding the direction of the causality in the coherence premium relationship. Although many sources indicate the direction, it is recommended to investigate this.

8.4.2 Data and Textual Analysis

The sample size should be extended as the required data is available. Rewriting the script to include the MD&A's from small business 10K's would add an interesting dimension and adds a representation of smaller firms. The cleaning process for the preparation of the data for analysis can be performed more precisely, for instance by considering the specific companies instead of following more general cut-off conditions. This will decrease the loss of valuable data and by looking into a wider range of outliers the amount of unreliable data with high leverage and influence can be decreased.
The accuracy of the premium measure data could be improved by using EBITDA instead of EBIT. Since depreciation and amortisation do not indicate real money spent, this would improve the accuracy of the ROA and EBIT margin premium measures. Additionally, different or more premium measures could be used.

The potential of textual analysis is far greater than has been shown thus far. Recommended extensions include:

- Instead of using single keywords, combinations can be counted as well, e.g. 'business sense' 'customer insight', 'product portfolio' etc.
- The difference in general occurrence of keywords could be taken into account by adding a factor to the different keywords instead of using logarithms in counting.
- Another option is the use of dequalifiers for keywords. To give an example, 'target' could be a keyword, but when it is used in the context of an acquisition target this might be very different. Using merger and acquisition as dequalifiers could exclude those occurrences from the keyword hit count. This could potentially reduce the number of false positives tremendously.
- The MD&A extraction script was not perfect and can possibly be further improved.
- Speed of execution improvements would be beneficial as well.
- It could also be helpful to exclude tables from the MD&A.
- Finally, instead of the MD&A other sources could be used for textual analysis as well.

8.4.3 Methodology

In further research on this topic the steps described in the methodology section should all gone through again. This is necessary because the options considered in this thesis are not exhaustive and because the test set that was used to make the decisions is limited in size. This research project can be used as inspiration to come up with more options. This begins by using a larger keyword longlist, based on more literature sources and more extensively enlarged with synonyms and similar words. Additionally, families of words can be added, e.g. adding strategically and strategic based on the keyword strategy.

An important step in the methodology derivation was the coding of the five CEO-letters. An important assumption that was made was that the CEO-letter is a decent proxy for the MD&A. A step of

uncertainty can be evaded here by using MD&A's instead of CEO-letters. Although the content is similar, the public, the level of aggregation and language used is at least somewhat different. A disadvantage is that the MD&A is much larger than the CEO-letter from the year report.
New research should improve this important step by using multiple coders and one controlling coder, who checks the coding of the 'normal' coders. In this way the analysis can be extended to more CEO-letters and the coding is not executed entirely by one person. The controlling coder ensures that the coding does not differ much between the normal coders. This would also result in a more extensive keyword database. This would enable strict and more robust selection of keywords without risking the final keyword list to become very small and thereby increasing the chances of obtaining many MD&A's with a zero coherence score. Additionally, the coding system of relevant/irrelevant and scoring 1-3 could be altered.
These changes will improve the process of translating the keyword longlist into a keyword shortlist. It will be then more extended and more precise than the current one. All the options considered in the methodology chapter and more can be used to find a better operationalization of coherence. It might even be possible to look into and optimise amongst all possible combinations of options by using modelling software.

Considering the test set used in the methodology derivation, a general improvement would be to use a different and larger one. Additionally, it could be extended to more industries and/or smaller companies could be added.

8.4.4 Future Research

This research has led to evidence that coherence generally leads to a positive difference in the four financial performance measures considered. Interesting next steps are to look into how much difference it makes and under what circumstances it makes a difference. The latter step was performed for several factors of influence already, but this should to be extended. Another improvement would be to improve the coherence score data/measurement to decrease the very large skew. This would enable the development of a regression model using the continuous coherence score data, instead of the ANOVA in which two groups are compared. Other data (from other industries) can be used as well.

Another interesting extension of this research is to look at the year to year differences in coherence and premium. This would answer the question whether a difference in coherence score is positively related with a difference in premium and could help to reveal the direction of the causal relationship.

Looking into other, possibly qualitative, measures for premium is interesting as well. Examples could be total shareholder return, earnings per share, economic value added, firm survival, firm reputation or customer satisfaction. The fact that large and diversified firms are more likely to have high coherence scores is an early indicator of a positive relationship between coherence and firm survival.

A more practical extension is to develop a tool to measure or rate the coherence in an individual company. A more advanced, extensive and/or reliable tool should be developed for this purpose, because the current measurement is based on a single data source and not accurate and reliable enough to draw hard conclusions from. The combination of an improved version of the current measurement with a more advanced version of the coherence survey from Booz & Company appears to be a promising option.

Bibliography

Acur, N., Kandemir D. and Boer, H. (2012), Strategic Alignment and New Product Development: Drivers and Performance Effects, *Journal of Product Innovation Management, Vol. 29, No. 2, pages 304–318.*

Architecting Value, a practitioner's perspective on IT strategy and architecture, accessed on June 15 2012 at *http://www.architectingvalue.com/science/*.

Barnett, W.P. and Burgelman, R.A. (1996), Evolutionary perspectives on strategy, *Strategic Management Journal, Vol. 17, pages 5-19.*

Barney, J. (1991), Firm resources and sustained competitive advantage, *Journal of Management, Vol. 17, No. 1, pages 99-120.*

Bowman, E.H. (1984), Content Analysis of Annual Reports for Corporate Strategy and Risk, *Interfaces, Vol. 14, No. 1,pages 61-71.*

Brown, S.V. and Tucker, J.W.U. (2011), Large-Sample Evidence on Firms' Year-over-Year MD&A Modifications, *Journal of Accounting Research, Vol. 40, No. 2, pages 309-346.*

Buzzell, R.D. and Gale, B.T. (1987), *The PIMS Principles: Linking Strategy to Performance*, Collier Macmillan, New York.

Campa, J.M. and Kedia, S. (2002), Explaining the diversification discount, *Journal of Finance, Vol. 57, No. 4, pages 1731-1762.*

Campbell-Hunt, C. (2000), What have we learned about Generic Competitive Strategy? A Meta-Analysis, *Strategic Management Journal, Vol. 21, No. 2, pages 127–154.*

Capon, C.; Farley, J.U. and Hoenig, S. (1990), Determinants of Financial Performance: A Meta-Analysis, *Management Science, Vol. 36, No. 10, Focused Issue on the State of the Art in Theory and Method in Strategy Research, pages 1143-1159.*

Chakravarthy, B.S. (1986), Measuring Strategic Performance, *Strategic Management Journal, Vol. 7, No. 5 , pages 437-458.*

Chmielewski, D.A. and Paladino, A. (2007), Driving a Resource Orientation: Reviewing the Role of Resource and Capability Characteristics, *Management Decision Vol. 45, No. 3, pages 462 – 483.*

Choe, C. and Yin, X. (2009), Diversification Discount, Information Rents, and Internal Capital Markets, *Quarterly Review of Economics and Finance, Vol. 49, No. 2, pages 178-196.*

Chorn, N.H. (1991), The "Alignment" Theory: Creating Strategic Fit, *Management Decision, Vol. 29, No. 1, pages 20-24.*

Derfus, P.J., Maggitti P.G., et al. (2008), The Red Queen effect: Competitive actions and firm performance, *The Academy of Management Journal, Vol. 51, No. 1, pages 61-80.*

Desarbo, W.S.; Di Benedetto, C.A. ; Song, M. and Sinha, I. (2005), Revisiting the Miles and Snow Strategic Framework: Uncovering Interrelationships between Strategic Types, Capabilities, Environmental Uncertainty, and Firm Performance, *Strategic Management Journal, Vol. 26, No. 1, pages 47-74*

Drazin, R. and Van de Ven, A.H. (1985), Alternative Forms of Fit in Contingency Theory, *Administrative Science Quarterly, Vol. 30, No. 4, pages 514-539.*

Eisenhardt, K.M. and Martin, J.A. (2000), Dynamic Capabilities: What are They?, *Strategic Management Journal, Vol. 21, No. 10/1, Special Issue: The Evolution of Firm Capabilities, pages 1105–1121.*

Farjoun, M. (2002), Towards an Organic Perspective on Strategy, *Strategic Management Journal, Vol. 23, No. 7, pages 561-594.*

Ginsberg, A. and Venkatraman, N. (1985), Perspectives of Organizational Strategy: A Critical Review of the Empirical Research, *The Academy of Management Review, Vol. 10, No. 3, pages 421-434.*

Grant, R.M. (1991), 'Contemporary Strategy Analysis', *Butler & Tanner LTD*, Frome and London.

Hall, R. (1993), A Framework Linking Intangible Resources and Capabilities to Sustainable Competitive Advantage, *Strategic Management Journal, Vol. 14, No. 8, pages 607-618.*

Hansen, G.S. and Wernerfelt, B. (1989), Determinants of Firm Performance: The Relative Importance of Economic and Organizational Factors, *Strategic Management Journal, Vol. 10, No. 5, pages 399-411.*

Hart, S. and Banbury, C. (1994), How Strategy-Making Processes Can Make a Difference, *Strategic Management Journal, Vol. 15, No. 4, pages 251-269.*

Impink, J. and Reppenhagen, D. (2012), Segments vs Subsidiaries as a Measure of Diversification, *unpublished.*

Jauch, L.R., Osborn, R.N. and Martin, T.N. (1980), Structured Content Analysis of Cases: A Complementary Method for Organizational Research, *Academy of Management Review, Vol. 5, No. 4, pages 517-525.*

Jick, T.D. (1979), Mixing Qualitative and Quantitative Methods: Triangulation in Action, *Administrative Science Quarterly, Vol. 24, No. 4, pages 602-611.*

Koller, T., Dobbs, R. and Huvett, B. (2010) *Value: The Four Cornerstones of Corporate Finance,* John Wiley & Sons

Laerd Statistics, provider of statistical help with SPSS, accessed through statistics.laerd.com/spss-tutorials in July and August 2012.

Leinwand, P. and Mainardi, C. (2010), The Coherence Premium, *Harvard Business Review, Vol. 88, No. 6, pages 86-92.*

Leinwand, P. and Mainardi, C. (2011a), *The Essential Advantage*, Harvard Business Review Press, 2011.

Leinwand, P. and Mainardi, C. (2011b), Sustainable Success: Winning Through Capabilities, *Business Strategy Review Vol. 22, No.2, pages 64–67.*

Long, C. and Vickers-Koch, M. (1995), Using Core Capabilities to Create Competitive Advantage, *Organizational Dynamics, Vol. 24, No. 1, pages 7–22.*

Markides, C.C. and Williamson, P.J. (1994), Related Diversification, Core Competencies and Corporate Performance, *Strategic Management Journal, Vol. 15, Special Issue: Strategy: Search for New Paradigms, pages 149-165.*

McArthur, A.W. and Nystrom, P.C. (1991), Environmental Dynamism, Complexity and Munificence as Moderators of Strategy-Performance Relationships*, Journal of Business Research, Vol. 23, pages 349-361 .*

McGahan, A.M. and Porter, M.E. (1997), How Much Does Industry Matter, Really?, *Strategic Management Journal, Vol. 18, Special Issue: Organizational and Competitive Interactions, pages 15-30.*

Miles, R.E. and Snow, C.C. (1978), *Organizational Strategy, Structure, and Process*, New York: McGraw-Hill.

Miller, D. (1992), Environmental Fit versus Internal Fit, *Organisation Science, Vol. 3, No. 2, pages 159-178.*

Miller, D. and Friesen, P.H. (1978), Archetypes of Strategy Formulation, *Management Science, Vol. 24, No. 9, pages 921-933.*

Miller, D. and Friesen, P.H. (1983), Strategy-Making and Environment: The Third Link, *Strategic Management Journal, Vol. 4, No. 3, pages 221-235.*

Naman, J.L. and Slevin, D.P. (1993), Entrepreneurship and the Concept of Fit: A Model and Empirical Tests, *Strategic Management Journal, Vol. 14, No. 2, pages 137-153.*

Newbert, S.L. (2007), Empirical Research on the Resource-Based View of the Firm: An Assessment and Sugestions for Future Research, *Strategic Management Journal, Vol. 28, pages 121-146.*

Newbert, S.L. (2008), Value, Rareness, Competitive Advantage, and Performance: A Conceptual-level Empirical Investigation of the Recourse-based View of the Firm, *Strategic Management Journal, Vol. 29, pages 745-768.*

Olson, E.M., Slater, S.F. and Hult, T.M. (2005), The Performance Implications of Fit Among Business Strategy, Marketing Organization Structure, and Strategic Behavior, *Journal of Marketing, Vol. 69, pages 49–65.*

Porter, M.E. (1979), How Competitive Forces Shape Strategy, *Harvard Business Review, Vol. 57, No. 2, pages 137-145.*

Porter, M.E. (1980), *Competitive Strategy*, Free Press, New York.

Porter, M.E. (1985), *Competitive Advantage*, Free Press, New York.

Powell, T.C. (1992), Organizational Alignment as Competitive Advantage, *Strategic Management Journal, Vol. 13, No. 2, pages 119-134.*

Prahalad, C.K. and Hamel, G. (1990), The Core Competence of the Corporation, *Harvard Business Review, Vol. 68, No. 3, pages 79–91.*

Prescott, J.E. (1986), Environments as Moderators of the Relationship between Strategy and Performance, *The Academy of Management Journal, Vol. 29, No. 2, pages 329-346.*

Ruefli, T.W. and Wiggins, R.R. (2003), Industry, Corporate, and Segment Effects and Business Performance: A Non-parametric Approach, *Strategic Management Journal, Vol. 24, No. 9, pages 861-879.*

Rumelt, R.P. (1991), How Much Does Industry Matter?, *Strategic Management Journal, Vol. 12, No. 3, pages 167-185.*

Servaes, H. (1996), The value of diversification during the conglomerate merger wave, *Journal of Finance, Vol. 51, No. 4, pages 1201-1225.*

Slater, S.F., Olson, E.M. and Hult, T.M. (2006), The Moderating Influence of Strategic Orientation on the Strategy Formation Capability-Performance Relationship, *Strategic Management Journal, Vol. 27, No. 12, pages 1221-1231.*

Teece, D.J. (2007), Explicating Dynamic Capabilities: The Nature and Microfoundations of (Sustainable) Enterprise Performance, *Strategic Management Journal, Vol. 28, pages 1319-1350.*

Treacy, M. And Wiersema, F. (1993), Customer Intimacy and Other Value Disciplines, *Harvard Business Review, Vol. 70, No. 1, pages 84-93.*

Venkatraman, N. and Camillus, J.T. (1984), Exploring the Concept of "Fit" in Strategic Management, *The Academy of Management Review, Vol. 9, No. 3, pages 513-525.*

Venkatraman, N. (1989a), Strategic Orientation of Business Entreprises: The Construct, Dimensionality and Measurement, *Management Science, Vol. 35, No. 8, pages 942-962.*

Venkatraman, N. (1989b), The Concept of Fit in Strategy Research: Toward Verbal and Statistical Correspondence, *The Academy of Management Review, Vol. 14, No. 3, pages 423-444.*

Venkatraman, N. and Prescott, J.E. (1990), Environment-Strategy Coalignment: An Empirical Test of Its Performance Implications, *Strategic Management Journal, Vol. 11, No. 1, pages 1-23.*

Varaiya, N.; Kerin, R.A. and Weeks, D. (1987), The Relationship between Growth, Profitability, and Firm Value, *Strategic Management Journal, Vol. 8, No. 5, pages 487-497.*

Villalonga, B. (2004) , Does Diversification Cause the "Diversification Discount"?, *Financial Management, Vol. 33, No. 2, pages 5-27.*

Vorhies, D.W. and Morgan, N.A. (2003), A Configuration Theory Assessment of Marketing Organization Fit with Business Strategy and Its Relationship with Marketing Performance, *Journal of Marketing, Vol. 67, No. 1, pages 100-115.*

Wernerfelt, B. and Montgomery C.A. (1988), Tobin's q and the Importance of Focus in Firm Performance, *The American Economic Review, Vol. 78, No. 1, pages 246-250.*

White, R.E. (1986), Generic Business Strategies, Organizational Context and Performance: An Empirical Investigation, *Strategic Management Journal, Vol. 7, No. 3, pages 217-231.*

Young, Ph. and Basarab, D. (2009), Business Acumen: Your Key to Success, *Learning Bursts Workbook, P&D Learning, LLC.*

Zegveld, M.A. (2004), Corporate Strategy and The Position of Technology: A Bird's Eye View, *Reader Technology and Strategy, Delft University of Technology*, ed. Zegveld, M.A. and Den Hartigh, E. *, pages 1–20.*

Zott, C. and Amit, R. (2008), The Fit Between Product Market Strategy and Business Model: Implications for Firm Performance, *Strategic Management Journal , Vol. 29, pages 1–26.*

12Manage, All you need to know about management, accessed on July 29 2012 at *http://www.12manage.com/methods_porter_value_chain_nl.html*

Appendices

I. Relevant Quantitative Literature Review

For the (Quantitative) Literature Review the reader is kindly referred to the file ***Relevant (Quantitative) Literature Review.xls***, to which access can be requested via the author.

II. Test Set

'The Coherence Premium' (Leinwand and Mainardi 2010) is the article that lays the foundation for this research project. Therefore this article was used as a means to test different methodological options and to decide on which path to pursue. The article substantiates the coherence premium with a graph of EBIT margin versus coherence score for seventeen companies in the period 2003-2007, which is shown in Figure 25 For this set a total of 36 MD&A's was found in the 10-K filings in the EDGAR database. The test set comprised of eight companies: Campbell's (2004-2007), Coca Cola (2003-2007), ConAgra (2005-2007), General Mills (2003-2007), Heinz (2003-2007), Kimberly-Clark (2003-2007), Kraft (2004-2007) and PepsiCo (2003-2007). Premium measures were obtained from EDGAR.

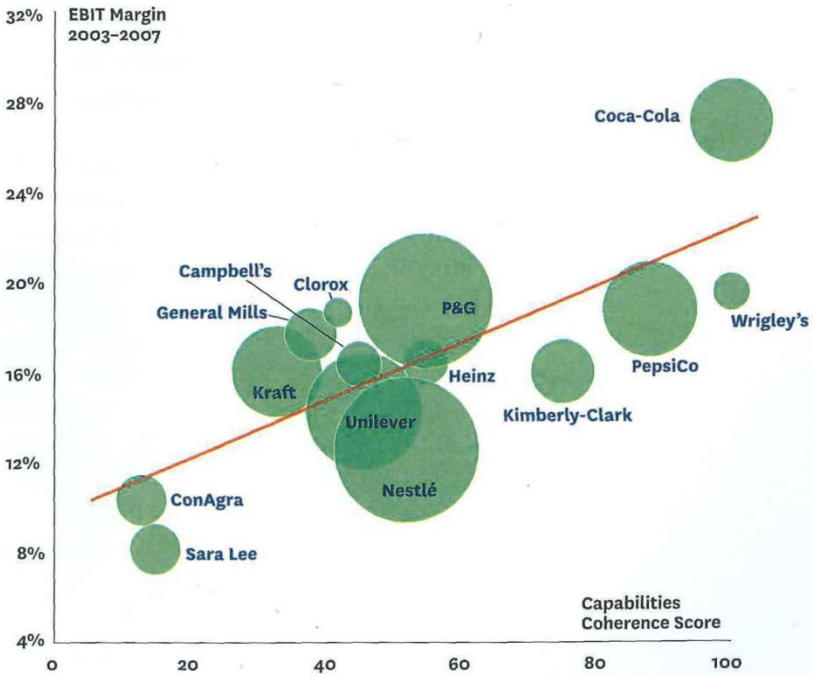

Figure 25 2003-2007 average EBIT margin versus coherence score for 17 consumer and retail companies. Reproduced from Leinwand and Mainardi, 2010, page 91.

Two analyses were performed in MS Excel 2010 to decide upon the final coherence measure. First the base case was tested on all the options to see whether alternative options would lead to better results.

The base case incorporated the following choices:

- General words only
- Include keywords with one or more hits

- Aggregated form of keywords
- Cutoff at 70% and average score of two
- Coherence construct from general model (multiplied elements)
- Normal counting per keyword
- Compare coherence and results from the same year

In varying the options it was found that improved R^2 values were found for:

- Aggregated plus form of keywords
- Cutoff at 70% and average score of 2.5
- Logarithmic counting per keyword: log(keyword hits*2)

With the combination of the better individual options an improved base case was created, which was again tested on improvement possibilities by changing the options individually. With one exception: The cut-off limit was lowered to an average score of two points as many words were lost in the step from two to 2.5 points on average.

In order to make sure the (local) optimum was found, the procedure was performed again with the improved case as base case for the choices that were not very clear (no significant R^2 differences between two or more viable options) from the previous base case exercise:

- Include keywords with at least zero, one or two hits
- Cut-off percentage and score
- Log, normal or square root counting per keyword
- Compare coherence to results from the same year

As this appendix was approximately 25 pages long and gave a very limited overview it is omitted here. The reader is referred to the Excel files ***Base Case Comparison.xls*** and ***Improved Case Comparison.xls***, to which access can be requested via the author. These files give a much better overview and therewith enable an effective comparison between the options. The P/E ratio and the ROA, EBITmargin and sales growth will be shown for each of the options inspected. It starts with the base case testing for one or more alternatives for the seven options initially chosen. Then the improved case is shown and subsequently compared to the alternatives of the four options that were not very distinctive in the base case comparison.

III. Keyword Lists

In this appendix the keyword longlist and the final keyword list are presented. The manual coding which led to the transformation of the keyword longlist into the final keyword list is too extended to add as an appendix. It can be found in the file Full dataset keywords.xls.

A. Keyword Longlist

General words

Element	Keywords		
Core Capabilities System	abilities	develop	learn, learns
	ability	developed	learned
	best practice	developing	learning
	build	development	practices
	building	differentiate, differentiates,	reinforce
	built	differentiated,	reinforced
	capabilities	differentiation	reinforces, reinforcing

	capability	distinct, distinctive	reinvented
	capable, capableness	enable	reinventing
	capacity, capacities	enabled	reinvention
	competence, competency,	enabler	resource
	competent, competencies,	enables	resources
	competences	enabling	science
	create	expert	scientific
	created	expertise	strength
	creates	experts	strengths
	creating	fortify, fortified, fortifying,	tool
	deepen, deepened, deepens	fortifies	tools
	deepening	knowledge	
Product & service portfolio	across	portfolio	products
	broad	divisions	scope
	broader	geographic	segment
	businesses	geographical	segments
	categories	geography	services
	diversify, diversified, diversifies,	industries	shared
	diversification	industry	sharing
	narrow	markets	throughout
Way to play	above-plan	contributed	
	actions	contributing	key
	add	contributions	method, methodology
	added	core	methods
	adding	couple, couples, coupling	mission
	addition	coupled	path
	additional	course	plan
	additionally	create	planned
	adds	created	planning
	align	creates	plans
	aligned	creating	position
	aligning	differentiate, differentiates,	positioned
	alignment	differentiated,	positioning
	approach	differentiation	positions
	attitude	direction	principal
	balance	distinguish, distinguishing,	priorities
	balanced	distinguished, distinguishes	priority
	best-positioned	emerge	refocused
	business	emerged	reinvent, reinvents
	capitalize, capitalized	emerges	reinvigorated
	capitalizing	establish	relevant
	chief	fit	reposition, repositions
	collaborate, collaborating	focus	repositioned
	collaborated	focused	repositioning
	collaboration	focusing	route
	competence, competency,	foundation	sense
	competent, competencies,	foundational	strategic
	competences	fundamental	understanding
	competitive	fundamentals	utmost
	competitors	strategically	value
	concentrate	strategies	value-added
	concentrated	strategy	values
	concentrates, concentrating	understand, understands,	vision
	contribute, contributes,	understood	way
	contribution	integral	ways

Specific words

Element	Aspect	Keywords		
Core Capabilities System	Financial management	evaluate, evaluates, evaluating, evaluated, evaluation	merge, merges, merger, merging, merged	performing
		financial	metric	profitability
		goal	metrics	profitable
		goals	objective	repurchase
		high-performing	objectives	result
		indicator	payback	results
		KPI	perform	ROE, ROA,ROC, ROS
		margin	performance	target
		margins	performed	targets
	Manufacturing	engineer, engineers, engineered, engineeres, engineering	plant	productiveness
		manufacturer	plants	productively
		manufacturers, manufactured,	produce	productivity
		manufactures, manufacture	produced producing	products
		manufacturing	produces	research
		package	product	technologies
		packaging	production	technology
			productive	technology-enabled
	Marketing	advertise, advertised	customers	promotions
		advertising	customers'	rebrand, rebranding,
		brand	go-to-market	rebrands
		branding	marketing	rebranded
		brands	promoting	route-to-market,
		consumer	promotion	routes-to-market
		consumers	promotional	speed to market
		customer		
	Portfolio Management	acquisition	divestiture,	investments
		acquisitions	divestitures,	portfolio
		add	divestment	project
		added	divestments	projects
		adding	exit, exited, exits,	reinvest
		addition	exiting	reinvested
		additional	invest	reinvesting
		additionally	investigating	
		adds	investing	
		divest, divests, divested, divesting,	investment	
	Sales	channel	demographic	sales
		consumer	demography	sell
		consumers	merchandise	selling
		customer	merchandising	sells
		customers	pricing	sold
		customers'		
	Supply Chain	channel	procurement	supply
		distributed, distribute, distributing,	relationship	transport, transports,
		distributes	relationships	transporting,
		distribution	served	transported
		inventory	service	truck
		lean	serving	vendor
		logistic, logistical	sourcing	vendors
		logistics	stock	warehouse,
		operations	supplier	warehouses

		planning	suppliers	warehousing
Way to play	Analyser	adapt adapted, adapting, adaptive analyse, analysed, analysing, analytics analysis analytical control coordinate	coordination, coordinating, coordinated, coordinates decision decisions	fine-tune information monitor system systems
	Defender	authorisation authorize authorized centralize, centralization centralized centralizing consolidate, consolidating, consolidator consolidated continual continually continue continued continues continuing cost costs effective effectively effectiveness efficiencies efficiency efficient execute executed, executes executing execution executive govern	identified, identifies, identifying image keep maintain, maintains, maintaining maintained niche optimize, optimized, optimization, optimizes optimizing price procedure, procedures, procedural proceed proceeding, proceeded proceeds process processes protect protected protects, protecting, protection quality remain remained remaining remains reputation	repute risktaker, risk-taker risk-takers risktaking saving savings scale secure, secures, securing, secured sizable size stabilize, stabilizes, stabilized, stabilization stable standard, standards, standardize, standardizes, standardization, standardized stay, stays, stayed, staying stick sticking sticks, sticked sticky value values volume
	Prospector	attractive, attractivity change changed changes changing customize, customizer, customization, customized, customizes decentralize, decentralizing decentralized distinguish, distinguishing, distinguished, distinguishes emerging enter, enters entered entering entrepreneur, entrepreneurs, entrepreneuring, entrepreneurship,	expanded expanding expands expansion explore explored, explores, exploring, exploration, explorative grow growing growth identification identify initiative initiatives innovation innovations innovative	introduce, introduces, introducing introduced introduction introductions launch, launching, launches launched lead leader leaders leadership leading market marketplace markets new novel

	entrepreneurialism, entrepreneurism entrepreneurial expand	innovator, innovated, innovativeness, innovates, innovated	opportunities opportunity test tested testing
Porter's Generic Strategies	differentiate, differentiates, differentiated, differentiation innovation innovations innovative	innovator, innovated, innovativeness, innovates, innovated price	segmentation, segmented, segmentations, segmenting
Leinwand and Mainardi Puretones	access, accessed, accessing aggregate, aggregator authorities bundle, bundles, bundling, bundled combinate, combining, combinating, combined, combinated, combination combine consolidate, consolidating, consolidator consolidated convenience convenient customize, customizer, customization, customized, customizes Disintermediate, disintermediates, disintermediating, disintermediated, disintermediator dominate, dominated, dominating, dominates, dominant durable, durability elegance, elegant engage, engages, engaged, engaging engagement engaging experience experiences government high-end influence	influencing, influenced, influences infrastructure innovation innovations innovative innovator, innovated, innovativeness, innovates, innovated insight intelligence introduce, introduces, introducing introduced introduction introductions leverage leveraged leverages leveraging merge, merges, merger, merging, merged mitigate, mitigated, mitigating, mitigation oversight platform pool, pooling, pooled, pools premium prevent, prevents,	prevented preventing price privilege, privileges, privileging, privileged regulate, regulates, regulated, regulating, regulation, regulation navigator, regulatory reputation risk, risks, risky, risk absorber rules solution, solutions provider solutions status system systems tailor, tailored, tailoring trust uncertainty unify, unifies, unifying, unification, unified, unite, united, uniting, unites value values
Treacy and Wiersema	closeness intimacy operations	personalize, personalization, personalized	product proximity
Venkatraman dimensions of strategy	aggressiveness aggressive defend, defends, defending, defensive, defensiveness, defended	forceful future futures futurity	proactive, proactivity, proactiveness riskiness

B. Final Keyword List

Element	Keywords		
Core Capabilities System	build building built	developed developing development	learns tool tools

	capabilities	enable	reinvented
	capability	enabled	reinventing
	deepen	enables	reinvention
	deepened	enabling	reinvent
	deepening	enabler	reinvents
	deepens	experts	resource
	knowledge	expertise	resources
	practices	experts	strength
	science	learned	strengths
	scientific	learning	
	develop	learn	
Product and Service Portfolio	geographic	throughout	categories
	geographical	segments	divisions
	geography	shared	across
	narrow	sharing	markets
	scope	portfolio	businesses
Way to Play	above-plan	core	principal
	actions	coupled	priority
	approach	couples	priorities
	attitude	couple	refocused
	align	coupling	refocuses
	aligned	fit	refocus
	aligning	focus	refocusing
	aligns	focused	relevant
	balanced	focusing	repositioned
	balance	foundation	repositioning
	alignment	foundational	reposition
	best-positioned	fundamental	repositions
	business	fundamentals	sense
	capitalizing	key	strategic
	capitalize	methods	strategies
	capitalized	method	strategy
	collaborated	methodology	strategically
	collaboration	mission	understand
	collaborate	path	understands
	collaborating	plan	understood
	competitive	planned	understanding
	competitors	planning	value
	competition	plans	value-added
	compete	positioned	vision
	competitor	positions	way
	competing	position	ways
	competed	positioning	

IV. Perl Scripts

In this appendix the scripts, programmed in Perl, that were used for the textual analysis are displayed and briefly explained. All these scripts are written by Joost Impink.

A. 10-K Download Script

The following script was used to download the 100.665 10-K SEC filings from the EDGAR database. For this purpose a text file was created in which all 10-K SEC-filings were given a unique id coupled to the url for the 10-K at the EDGAR database (list available from EDGAR).

```
#!/usr/bin/Perl
use LWP;
use HTTP::Request;
sub get_http
{
	my $url = shift;
	my $request = HTTP::Request->new(GET => $url);
	my $response = $ua->request($request);
	if (!$response->is_success)
	{
		print STDERR "GET '%s' failed: %s\n", $url,
$response->status_line;
		return undef;
	}
	return $response->content;
}
# user agent object for handling HTTP requests
my $ua = LWP::UserAgent->new;

# if you only want a portion of the filing, un-comment the next line
#$ua->max_size(50000);  # 50k byte limit

########################## write dir , use "\\" and not "\", for example: "C:\\temp"
$write_dir = "C:/Perl/Download";
########################## write dir
########################## filename with urls (put in same directory as script)
open dlthis, "c_10K_list.txt" or die $!;
########################## filename with urls (put in same directory as script)
########################## log
open LOG , ">download_log.txt" or die $!;
########################## log

my @file = <dlthis>;

foreach $line (@file) {
#CIK, filename, blank is not used (included because it will capture the newline)
	($CIK, $get_file, $blank) = split (",", $line);
	$get_file = "http://www.sec.gov/Archives/" . $get_file;
	$_ = $get_file;

if ( /([0-9|-]+).txt/ ) {
	$filename = $write_dir . "/" . $CIK . ".txt";
	open OUT, ">$filename" or die $!;
	print "file $CIK \n";
	my $request = HTTP::Request->new(GET => $get_file);
	my $response =$ua->get($get_file );	$p = $response->content;
	if ($p) {
		print OUT $p;
		close OUT;
		} else {		#error logging
		print LOG "error in $filename - $CIK \n" ;
		}
	}
}
close LOG; #ignore the line below (inserted by Forum engine because it wants to 'close' a similar tag used to load the file)
```

B. MD&A Extraction Script

The following script was used to extract Item 7, Management Discussion and Analysis, from the 10-K filings. The script first filters the html code with the stripscript package and writes the clean 10-K to a new text file. Then it identifies the beginning and the end of item 7 and writes Item 7 to a new text file. This process is not straightforward as not all K-10's have an MD&A and Item 7 (and Item 7A and 8, where the MD&A ends) occurs frequently in tables of contents and internal referencing. Many filtering

approaches were investigated. In the end the best results were obtained with demanding a minimal size of 10.000 characters for the MD&A and using certain dequalifiers. The minimal size excludes tables of contents and MD&A's by reference, in which the reader is referred to another document for the MD&A. The dequalifiers signal the use of Item 7/7A/8 as a reference to the real Item 7/7A/8 paragraph. Dequalifiers included see, in, Item 7/7A/8 within a sentence and continued. Additionally the script counts the words, filtered for frequently occurring words like the, a, an, for, with, do etc. with the StopWords package. This output is written to another text file.

```
#!/usr/bin/Perl
use LWP;

use HTML::StripScripts;
use Lingua::Stem;
use Lingua::EN::StopWords qw(%StopWords);

use HTML::Restrict;

# on the command line (">" betekent stuur output naar file ipv scherm)
# ga met 'cd' in naar de juiste directory (waar Perl script staat), en dan:
# Perl scan_keywords_MD&A.pl > output_firm_year.csv
# zorg dat 10Ks in een subdirectory staan ($dirIn, hieronder)
# in dirOut worden dan files weggeschreven (de 10K zonder HTML, de MD&A, en
# een textfile met de woord-count)
# naar het scherm (of > file) wordt 1 regel met de lengte van de MD&A geprint
# als deze lengte vrij kort is (bijv. 1000 characters), dan gaat het waarschijnlijk
# om de table of contents
# (ik pak de 'laatste' item 7 en de laatste item 8, als het alleen in de table of contents
# staat (bijv. 'incuded by reference'), dan komt het niet terug in de text

 $dirIn = "C:/Perl/Data/Retail/";
$dirOut = "C:/Perl/Results/Retail_190712_1/";

 sub getMD&A {

        my $filing_case = shift;

        my $filing = lc($filing_case);

        my $tag7  = "item 7";
        my $tag8  = 'item 8';

        my @items7  = ();
        my @items7a = ();
        my @items8  = ();

        my $filing_copy = $filing;

        my $cumOccur = 0;

        #scan for "item"

        do {
                my $occurance = index($filing, "item");

                # figure out what item this is
                my $strTemp = substr($filing, $occurance-10, 160);
                my $strTempSmall = substr($filing, $occurance-10, 20);
                # 'see item',  'in item', 'and item', .. and 'continued' may be ignored .. ('scontinued' is fine -- discontinued operations)
                #print "found: --$strTempSmall-- \n";
                if ( ($strTempSmall =~ m/see\s*\n?\s*item/) ||
                        ($strTempSmall =~ m/part\s*\n*\s*ii\,?\s*\n*\s*item/) ||
                        ($strTempSmall =~ m/and\s*\n*\s*item/) ||
                       ($strTempSmall =~ m/in\s*\n*\s*item/) ||
```

```
                ($strTempSmall =~ m/item\s*\n?\s*6/) ||
                ($strTemp =~ m/[^s]continued/) ) {
        # next line: debugging info (matched string is saved)

        #if (($strTempSmall =~ m/(see\s*\n?\s*item)/) || ($strTempSmall =~ m/(part\s*\n*\s*ii\,?\s*\n*\s*item)/)|| ($strTempSmall =~
m/(and\s*\n*\s*item)/) || ($strTempSmall =~ m/(in\s*\n*\s*item)/) || ($strTemp =~ m/([^s]continued)/) ) {
                #ignore
                #print "BAD: $strTemp \n";
                #print "matched on --$1--\n";
        } else {

                #does it start on a new line? (is there a word in front?)
                #difficulty with doing this in one expression
                # first expression tests for leading whitespace (which is ok)
                # second one detects a word (which is not ok)
                # when using the first expression only, somehow some 'empty' trailing space is not matched (10K 3461 for example)

                # not happy with second part => in-text 'part ii, item 7, md&a' will make it as start of md&a (while it is not)
                #if (($strTemp =~ m/\n\s*item/i) || !($strTemp =~ m/\w\s*item/i)) {

                if ($strTemp =~ m/\n\s*item/i) {

                        # strip enters
                        $strTemp =~ s/\n//g;
                        $strTempSmall =~ s/\n//g;

                        # is it follwed by whitespace/period/whitespace and a 7 or an 8?
                        if (($strTempSmall =~ m/item\s*.?\s*[78]/i)) {

                                #yes, good; strip newlines (HTML may cause excessive use of newlines)

                                #print "GOOD: $strTemp \n";


                                # strip dash
                                $strTemp =~ s/-//g;

                                #item 7a
                                if ($strTemp =~ m/\s*item\s*\da/i) {

                                        if (($strTemp =~ m/quantitative\s*disclosure/) ||
                                                ($strTemp =~ m/qualitative\s*disclosure/) ){


                                                # item 7: Risk
                                                #print "Risk " . $strTemp . "\n";
                                                push(@items7a, $cumOccur + $occurance);

                                                # I am pushing it only items8 even though it is item7a
                                                # (to use single array to determine ending point)
                                                push(@items8, $cumOccur + $occurance);

                                        }
                                }

                                #item 7 or 8
                                if ($strTemp =~ m/\s*item\s*[78]/) {

                                        #print "item 7 or 8 \n";

                                        if ($strTemp =~ m/discussion\s*and\s*analysis/) {

                                                # item 7: MD&A
                                                #my $tempPos = $cumOccur + $occurance;
                                                #print "MD&A ($tempPos) " . $strTemp . "\n";
                                                push(@items7, $cumOccur + $occurance);
```

```
                                        } elsif ($strTemp =~ m/financial\s*statements/) {

                                                # item 8: Financial statements
                                                #print "Financial Statements " . $strTemp . "\n";
                                                push(@items8, $cumOccur + $occurance);
                                        }
                                }
                        }
                }

        }

        $filing = substr($filing, $occurance + 10);
        $cumOccur += $occurance +10;
}       while (index($filing, "item") > 0 );

#print "items 7: \n";
foreach $item(@items7) {
#               print " $item \n";
                if ($item > $max7) { $max7 = $item;}
}

#       print "\n\nitems 8: \n";
foreach $item(@items8) {
#               print " $item \n";
                if ($item > $max8) { $max8 = $item;}
}

my $start = shift(@items7);
my $end = shift(@items8);

do {
        if ($end > $start) {

                if ( ($end - $start) > 10000) {

                        #
                        #       it is possible that item 7 MD&A is found in toc, but not in full text
                        #       if so, @items7 will have a single entry
                        #       the 'next' item 7 will be false

                        if ($start) {
                                return substr($filing_copy, $start, $end - $start);
                        } else {
                                return;
                        }

                } else {

                        $start = shift(@items7);
                }
        }

} while ($end = shift(@items8));

# no valid hit
return;

#currently, the code below in this function is not executed

# determine start and end of MD&A

# keep the highest occurences
my $max7 = 0;
my $max7a = 0;
```

```
	my $max8 = 0;
	reset(@items7);
	reset(@items7a);
	reset(@items8);

	print "items 7: \n";
	foreach $item(@items7) {
		print " $item \n";
		if ($item > $max7) { $max7 = $item;}
	}

	print "items 7a: \n";
	foreach $item(@items7a) {
		print " $item \n";
		if ($item > $max7a) { $max7a = $item;}
	}

	print "\n\nitems 8: \n";
	foreach $item(@items8) {
		print " $item \n";
		if ($item > $max8) { $max8 = $item;}
	}

	reset(@items7);
	reset(@items7a);
	reset(@items8);

	# use risk factor paragraph if it exists
	# I use the max, because it is often incorporated by reference
	# (if that is the case, it will show in the table of contents)

	if ($max7a > 0) {

		#risk factors exist, is it just a table of contents entry (included by reference)
		if ($max7a < $max7) {

			#the last occurance of risk factors is before the md&a => ignore

		} else {

			#max7a is valid position for risk factors, use it as ending marker (end of MD&A)
			$max8 = $max7a;
		}
	}

	#print "\n max7: $max7 , max7a: $max7a , max8: $max8 \n";
	if (length(@items7) < 5 && $max8 > 0 && $max7 > 0) {

		return substr($filing_copy, $max7, $max8 - $max7);
	}

	return ;
}

opendir(DIR, $dirIn);

foreach my $file (readdir(DIR))   {

	$i++;
#	if ($i == 6) { last };

	if($file =~ m/txt/) {

		local( $/, *FH ) ;
		open( FH, $dirIn . $file ) or die "fatal error reading $file\n";
		$filing_raw = <FH>;
```

```
		# remove html tags
		my $hr = HTML::Restrict->new();
		my $str10K = $hr->process($filing_raw);

		# remove   etc (replace by space)
		$str10K =~ s!&[^;\s]+;! !g;

		open (MYFILE, '>' . $dirOut . "_10K" . $file );
		print MYFILE $str10K or die ("cannot write $file");
		close (MYFILE);

		$strMD&A = getMD&A($str10K);
		open (MYFILE, '>' . $dirOut . "_MD&A" . $file );
		print MYFILE $strMD&A or die ("cannot write $file");
		close (MYFILE);

		printf ("%s,%d\n", $file, length($strMD&A));

		#convert to array of lines
		@arrFile = split("\n",$strMD&A );

		my %count = ();

		# word count
		foreach $item (@arrFile) {
			$_ = $item;
			tr/A-Za-z/ /cs;
			++$count{$_} foreach split(' ', lc $_);
		}

		# build array with word count, ignoring stop words
		my @lines = ();
		my ($w, $c);
		while (($key, $value) = each(%count)){
			if (grep { !$StopWords{$_} } $key) {
				if (length($key) > 2) {
					push(@lines, sprintf("%s,%d\n", $key, $value));
				}
			}
		}

		#write output to file
		open (MYFILE, '>' . $dirOut . "_score" . $file );
		print MYFILE sort { $a cmp $b } @lines;
		close (MYFILE);
	}
}
```

C. Counting Keywords Script and Total Words Script

The previous script gave the word count in the MD&A as an output. The first script displayed here takes the word count for the relevant coherence keywords, which are defined by a list in an external .csv file, and exports them to a new .csv file. The second script calculates the total word hits in the MD&A, to give a measure for the size of the MD&A. This is used to relate the relevant keywords hits to the MD&A size in terms of total words.

Counting Keywords Script

```
#!/usr/bin/Perl
use LWP;

# dit is de directory met _score files
#------------------------------------------------------------
$dir = "C:/perl/results/AllDataNonzero/";
```

```
#-----------------------------------------------------------------

# maak voor elke metric een .csv of .txt file met de keywords (voorbeeld: zie score_metric1.csv)
# run dit script, Perl score_keywords_MD&A.pl > score_metric1.csv
# (voor elke metric een aparte file out)
# de output is de word count (een kolom voor elk keyword; excel kan 16,000 kolommen aan)
# een rij per filing

# open file with scoring keywords

#-----------------------------------------------------------------
 open(FILENAME,"<Element_scores.csv");
#-----------------------------------------------------------------
 $row = 0;
 my @word = ();
 print "id";
 while ( <FILENAME> )
 {
   chomp;
   $row +=1;
   @line = split /,/;
   push(@word, $line[0]);
   print "," . $line[0];
 }
 print "\n";
 close(FILENAME);

#second set of headers
print "id";
for ($j=1; $j <= $row; $j++) {

                print ",word$j";
}
print "\n";

$i = 0;
opendir(DIR, $dir);

foreach my $file (readdir(DIR))   {

                $i++;

                if($file =~ m/_score/) {

                                # clear score hash
                                %fileScore = ();

                                #load score file into array myScore

                                open(FILENAME2,$dir . $file);

                                $row = 0;
                                while ( <FILENAME2> )
                                 {
                                   chomp;
                                   @line = split /,/;
                                   $fileScore{$line[0]} = $line[1];
                                  # print "setting score " . $line[0] . " to " . $line[1] . "\n";
                                 }

                                close(FILENAME2);

                                #print file id (numbers preceding '.txt')
                                $file =~ m/(\d+)\.txt$/;
                                print $1;
```

```
                        #print score
                        foreach $keyword(@word) {

                                if ($val = $fileScore{$keyword}){
                                        print "," . $val;
                                        #print $keyword . " => " . $val . "\n";
                                } else {
                                        print ",0";
                                        #print $keyword . " => 0 \n";
                                }
                        }
                        print "\n";
                }
}

closedir(DIR);
```

Counting Total words Script

```
#!/usr/bin/Perl
use LWP;

# dit is de directory met _score files
#-----------------------------------------------------------------
$dir = "C:/perl/results/AllDataNonzero/";
#-----------------------------------------------------------------

# maak voor elke metric een .csv of .txt file met de keywords (voorbeeld: zie score_metric1.csv)
# run dit script, Perl score_keywords_MD&A.pl > score_metric1.csv
# (voor elke metric een aparte file out)
# de output is de word count (een kolom voor elk keyword; excel kan 16,000 kolommen aan)
# een rij per filing

# open file with scoring keywords

#-----------------------------------------------------------------
 open(FILENAME,"<Element_scores.csv.csv");
#-----------------------------------------------------------------
$row = 0;
my @word = ();
print "id";
while ( <FILENAME> )
{
  chomp;
  $row +=1;
  @line = split /,/;
  push(@word, $line[0]);
  print "," . $line[0];
 }
print ",totalWords\n";
close(FILENAME);

#second set of headers
print "id";
for ($j=1; $j <= $row; $j++) {

        print ",word$j";
}
print ",totalWords\n";

$i = 0;
opendir(DIR, $dir);


foreach my $file (readdir(DIR))   {
```

```
	$i++;

	if($file =~ m/_score/) {

		# clear score hash
		%fileScore = ();

		#load score file into array myScore

		open(FILENAME2,$dir . $file);

		$row = 0;
		$totalWords = 0;
		while ( <FILENAME2> )
		{
		  chomp;
		  @line = split /,/;
		  $fileScore{$line[0]} = $line[1];
		 # print "setting score " . $line[0] . " to " . $line[1] . "\n";
		  $totalWords += $line[1];
		}

		close(FILENAME2);

		#print file id (numbers preceding '.txt')
		$file =~ m/(\d+)\.txt$/;
		print $1;

		#print score
		foreach $keyword(@word) {

			if ($val = $fileScore{$keyword}){
				print "," . $val;
				#print $keyword . " => " . $val . "\n";
			} else {
				print ",0";
				#print $keyword . " => 0 \n";
			}
		}
		print ",$totalWords\n";
	}
}

closedir(DIR);
```

V. Analysis of Variance and Kruskal-Wallis Results

In this appendix the full ANOVA results, consisting of the descriptives, test of homogeneity of variances and ANOVA details, are presented for the full large sample, for the individual factor correction and for the multiple factor correction analyses. Additionally the Kruskal Wallis Results are added to indicate the significance of the difference between the groups, when the assumption of equal variances is not met.

A. Full Large Sample

The full ANOVA and Kruskal-Wallis results for the full large sample are given here.

Descriptives

		N	Mean	Std. Deviation	Std. Error	95% Confidence Interval Mean		Min.	Max.
						Lower Bound	Upper Bound		
Sales growth	Low Coherence	2221	.09456	.302488	.006419	.08197	.10715	-1.000	3.365
	High Coherence	2221	.12087	.340822	.007232	.10669	.13505	-1.000	3.357
	Total	4442	.10772	.322458	.004838	.09823	.11720	-1.000	3.365
ROA	Low Coherence	2263	-.00805	.228061	.004794	-.01745	.00135	-3.026	2.372
	High Coherence	2288	.00292	.203315	.004251	-.00541	.01126	-2.630	1.091
	Total	4551	-.00253	.216021	.003202	-.00881	.00375	-3.026	2.372
EBIT margin	Low Coherence	2254	.04458	.083659	.001762	.04112	.04803	-.850	.460
	High Coherence	2248	.05025	.122222	.002578	.04519	.05530	-1.303	.692
	Total	4502	.04741	.104733	.001561	.04435	.05047	-1.303	.692
P/E ratio	Low Coherence	1271	23.102	34.9334	.979868	21.180	25.025	1.161	530.13
	High Coherence	1554	25.458	35.4776	.899971	23.693	27.224	1.218	485.88
	Total	2825	24.398	35.2471	.663153	23.098	25.699	1.161	530.13

Table 29 Descriptives of the full large sample.

ANOVA

		Sum of Squares	df	Mean Square	F	Sig.
sales growth	Between Groups	.769	1	.769	7.405	.007***
	Within Groups	461.002	4440	.104		
	Total	461.771	4441			
ROA	Between Groups	.137	1	.137	2.937	.087*
	Within Groups	212.188	4549	.047		
	Total	212.325	4550			
EBIT margin	Between Groups	.036	1	.036	3.299	.069*
	Within Groups	49.335	4500	.011		
	Total	49.371	4501			
P/E ratio	Between Groups	3881.657	1	3881.657	3.127	.077*
	Within Groups	3504528.852	2823	1241.420		
	Total	3508410.509	2824			

Table 30 ANOVA details for the full large sample. The suffixes *, ** and *** indicate that differences between the high and low coherence groups are significant at the 0.10, .05 and .01 significance level respectively.

Test of Homogeneity of Variances

	Levene Statistic	df1	df2	Sig.
Sales growth	1.574	1	4440	.210
ROA	3.355	1	4549	.067
EBIT margin	21.778	1	4500	.000*

P/E ratio	.942	1	2823	.332

Table 31 Levene's test of homogeneity of variances for the full large sample. Suffix * indicates that the variances between the two coherence groups are unequal at the 0.05 significance level. This means that the ANOVA assumption of equal variances is not met and this can influence the ANOVA results.

Independent Samples Kruskal-Wallis Test (Non-parametric One-way ANOVA)	
	Sig.
Sales growth	.001***
ROA	.000***
EBIT margin	.000***
P/E ratio	.000***

Table 32 Kruskal-Wallis Test (Non-parametric One-way ANOVA) for the full large sample. The suffixes *, ** and *** indicate that differences between the high and low coherence groups are significant at the 0.10, .05 and .01 significance level respectively.

B. Corrected for Relevant Factors Individually

The full ANOVA and Kruskal-Wallis results corrected for relevant factors individually are given here. Results are presented in the order size, diversification and industry (2-digit SIC code).

1. Size

The full ANOVA and Kruskal-Wallis results for the three size segments are given here.

Descriptives

			N	Mean	Std. Deviation	Std. Error	95% Confidence Interval for Mean		Minimum	Maximum
							Lower Bound	Upper Bound		
Small	**Sales growth**	Low Coherence	835	.08188	.361021	.012494	.05736	.10641	-1.000	3.282
		High Coherence	656	.13623	.425464	.016612	.10361	.16884	-1.000	3.249
		Total	1491	.10579	.391481	.010138	.08591	.12568	-1.000	3.282
	ROA	Low Coherence	856	-.05356	.304379	.010403	-.07398	-.03314	-3.026	2.372
		High Coherence	709	-.07001	.291193	.010936	-.09148	-.04854	-2.630	.570
		Total	1565	-.06101	.298495	.007545	-.07581	-.04621	-3.026	2.372
	EBIT margin	Low Coherence	847	.02004	.105545	.003627	.01292	.02716	-.850	.412
		High Coherence	669	-.00444	.184967	.007151	-.01848	.00960	-1.303	.692
		Total	1516	.00924	.146471	.003762	.00186	.01662	-1.303	.692
	P/E ratio	Low Coherence	402	21.567	31.4913	1.570645	18.480	24.655	1.300	372.68
		High Coherence	342	27.227	42.5642	2.301607	22.700	31.754	1.222	451.07
		Total	744	24.169	37.0766	1.359295	21.500	26.837	1.222	451.07
Medium	**Sales growth**	Low Coherence	763	.10316	.270802	.009804	.08392	.12241	-.783	3.365
		High Coherence	742	.13156	.328129	.012046	.10791	.15521	-.539	3.341
		Total	1505	.11717	.300671	.007750	.10196	.13237	-.783	3.365
	ROA	Low Coherence	780	.00939	.174154	.006236	-.00285	.02163	-1.570	1.655
		High Coherence	753	.01472	.159841	.005825	.00328	.02615	-1.539	1.091
		Total	1533	.01200	.167243	.004271	.00363	.02038	-1.570	1.655
	EBIT margin	Low Coherence	780	.06015	.069787	.002499	.05524	.06505	-.207	.460
		High Coherence	753	.06221	.077123	.002811	.05669	.06772	-.605	.372
		Total	1533	.06116	.073465	.001876	.05748	.06484	-.605	.460
	P/E ratio	Low Coherence	436	22.926	28.5649	1.368011	20.237	25.614	1.289	358.02
		High Coherence	517	27.440	40.3734	1.775619	23.952	30.929	1.322	485.88
		Total	953	25.375	35.5161	1.150479	23.117	27.633	1.289	485.88
Large	**Sales growth**	Low Coherence	623	.10101	.248244	.009946	.08148	.12054	-.618	2.446
		High Coherence	823	.09900	.267661	.009330	.08068	.11731	-.662	3.357
		Total	1446	.09986	.259388	.006821	.08648	.11324	-.662	3.357

	ROA	Low Coherence	627	.03240	.137150	.005477	.02164	.04315	-1.018	2.170
		High Coherence	826	.05478	.105562	.003673	.04757	.06199	-2.075	.344
		Total	1453	.04512	.120681	.003166	.03891	.05133	-2.075	2.170
	EBIT margin	Low Coherence	627	.05836	.052606	.002101	.05423	.06248	-.094	.362
		High Coherence	826	.08364	.062378	.002170	.07938	.08790	-.364	.380
		Total	1453	.07273	.059672	.001565	.06966	.07580	-.364	.380
	P/E ratio	Low Coherence	433	24.705	42.9024	2.061756	20.653	28.758	1.161	530.13
		High Coherence	695	23.114	26.5418	1.006788	21.137	25.091	1.218	374.42
		Total	1128	23.725	33.7631	1.005282	21.752	25.697	1.161	530.13

Table 33 Descriptives of the three size samples.

ANOVA

			Sum of Squares	df	Mean Square	F	Sig.
Small	**sales growth**	Between Groups	1.085	1	1.085	7.108	.008***
		Within Groups	227.268	1489	.153		
		Total	228.353	1490			
	ROA	Between Groups	.105	1	.105	1.178	.278
		Within Groups	139.246	1563	.089		
		Total	139.351	1564			
	EBIT margin	Between Groups	.224	1	.224	10.507	.001***
		Within Groups	32.278	1514	.021		
		Total	32.502	1515			
	P/E ratio	Between Groups	5918.874	1	5918.874	4.325	.038**
		Within Groups	1015465.926	742	1368.552		
		Total	1021384.800	743			
Medium	**sales growth**	Between Groups	.303	1	.303	3.361	.067*
		Within Groups	135.663	1503	.090		
		Total	135.966	1504			
	ROA	Between Groups	.011	1	.011	.389	.533
		Within Groups	42.840	1531	.028		
		Total	42.851	1532			
	EBIT margin	Between Groups	.002	1	.002	.301	.583
		Within Groups	8.267	1531	.005		
		Total	8.268	1532			
	P/E ratio	Between Groups	4821.072	1	4821.072	3.833	.051*
		Within Groups	1196024.874	951	1257.650		
		Total	1200845.947	952			
Large	**sales growth**	Between Groups	.001	1	.001	.021	.884
		Within Groups	97.221	1444	.067		
		Total	97.222	1445			
	ROA	Between Groups	.179	1	.179	12.358	.000***
		Within Groups	20.968	1451	.014		
		Total	21.147	1452			
	EBIT margin	Between Groups	.228	1	.228	66.874	.000***
		Within Groups	4.942	1451	.003		
		Total	5.170	1452			
	P/E ratio	Between Groups	675.419	1	675.419	.592	.442
		Within Groups	1284044.823	1126	1140.360		
		Total	1284720.242	1127			

Table 34 ANOVA details for the size segments. The suffixes *, ** and *** indicate that differences between the high and low coherence groups are significant at the 0.10, .05 and .01 significance level respectively.

Test of Homogeneity of Variances

		Levene Statistic	df1	df2	Sig.
Small	**Sales growth**	5.300	1	1489	.021*

	ROA	.377	1	1563	.539
	EBIT margin	38.132	1	1514	.000*
	P/E ratio	4.419	1	742	.036*
Medium	**Sales growth**	5.178	1	1503	.023*
	ROA	2.489	1	1531	.115
	EBIT margin	.038	1	1531	.846
	P/E ratio	8.593	1	951	.003*
Large	**Sales growth**	2.131	1	1444	.145
	ROA	3.612	1	1451	.058
	EBIT margin	16.373	1	1451	.000*
	P/E ratio	6.344	1	1126	.012*

Table 35 Levene's test of homogeneity of variances for the three size segment samples. Suffix * indicates that the variances between the two coherence groups are unequal at the 0.05 significance level. This means that the ANOVA assumption of equal variances is not met and this can influence the ANOVA results

Independent Samples Kruskal-Wallis Test (Non-parametric One-way ANOVA)		
		Sig.
Small	**Sales growth**	.001***
	ROA	.432
	EBIT margin	.421
	P/E ratio	.002***
Medium	**Sales growth**	.139
	ROA	.020**
	EBIT margin	.179
	P/E ratio	.358
Large	**Sales growth**	.900
	ROA	.000***
	EBIT margin	.000***
	P/E ratio	.004***

Table 36 Kruskal-Wallis Test (Non-parametric One-way ANOVA) for the full large sample. The suffixes *, ** and *** indicate that differences between the high and low coherence groups are significant at the 0.10, .05 and .01 significance level respectively.

2. Diversification

The full ANOVA and Kruskal-Wallis results for the diversified and undiversified groups are given here.

Descriptives

			N	Mean	Std. Deviation	Std. Error	95% Confidence Interval for Mean		Minimum	Maximum
							Lower Bound	Upper Bound		
Undiversified	**Sales growth**	Low Coherence	1766	.08675	.290276	.006907	.07320	.10029	-1.000	3.282
		High Coherence	1462	.13184	.359892	.009412	.11338	.15030	-1.000	3.341
		Total	3228	.10717	.324392	.005710	.09598	.11837	-1.000	3.341
	ROA	Low Coherence	1796	-.00845	.222551	.005251	-.01875	.00185	-2.470	2.372
		High Coherence	1505	-.00693	.232198	.005985	-.01867	.00481	-2.630	1.091
		Total	3301	-.00776	.226967	.003950	-.01550	-.00001	-2.630	2.372
	EBIT margin	Low Coherence	1793	.04227	.085979	.002031	.03829	.04625	-.850	.460
		High Coherence	1476	.04321	.131762	.003430	.03648	.04994	-1.278	.417
		Total	3269	.04269	.109040	.001907	.03895	.04643	-1.278	.460
	P/E ratio	Low Coherence	1015	22.455	31.9329	1.002318	20.488	24.421	1.161	530.13
		High Coherence	976	26.365	38.4297	1.230103	23.951	28.779	1.218	485.88
		Total	1991	24.371	35.3128	.791400	22.819	25.923	1.161	530.13
Diversi	**Sales**	Low Coherence	455	.12488	.344507	.016151	.09314	.15662	-.618	3.365

fied	**growth**	High Coherence	759	.09974	.299795	.010882	.07838	.12111	-.873	3.357
		Total	1214	.10917	.317385	.009109	.09129	.12704	-.873	3.365
	ROA	Low Coherence	467	-.00651	.248362	.011493	-.02909	.01607	-3.026	2.170
		High Coherence	783	.02187	.129036	.004611	.01282	.03093	-1.296	.273
		Total	1250	.01127	.183378	.005187	.00109	.02145	-3.026	2.170
	EBIT margin	Low Coherence	461	.05356	.073343	.003416	.04685	.06028	-.504	.257
		High Coherence	772	.06370	.100231	.003607	.05662	.07079	-1.303	.692
		Total	1233	.05991	.091212	.002598	.05482	.06501	-1.303	.692
	P/E ratio	Low Coherence	256	25.670	44.8847	2.805291	20.146	31.195	1.575	522.43
		High Coherence	578	23.928	29.8066	1.239789	21.493	26.363	1.464	374.42
		Total	834	24.463	35.1108	1.215787	22.076	26.849	1.464	522.43

Table 37 Descriptives of the undiversified and diversified samples.

ANOVA

numSegs (Binned)			**Sum of Squares**	**df**	**Mean Square**	**F**	**Sig.**
Undiversified	**Sales growth**	Between Groups	1.626	1	1.626	15.526	.000***
		Within Groups	337.951	3226	.105		
		Total	339.578	3227			
	ROA	Between Groups	.002	1	.002	.036	.849
		Within Groups	169.994	3299	.052		
		Total	169.996	3300			
	EBIT margin	Between Groups	.001	1	.001	.060	.806
		Within Groups	38.855	3267	.012		
		Total	38.856	3268			
	P/E ratio	Between Groups	7607.637	1	7607.637	6.116	.013**
		Within Groups	2473903.679	1989	1243.793		
		Total	2481511.316	1990			
Diversified	**Sales growth**	Between Groups	.180	1	.180	1.786	.182
		Within Groups	122.010	1212	.101		
		Total	122.189	1213			
	ROA	Between Groups	.236	1	.236	7.042	.008***
		Within Groups	41.765	1248	.033		
		Total	42.001	1249			
	EBIT margin	Between Groups	.030	1	.030	3.575	.059*
		Within Groups	10.220	1231	.008		
		Total	10.250	1232			
	P/E ratio	Between Groups	538.408	1	538.408	.436	.509
		Within Groups	1026355.879	832	1233.601		
		Total	1026894.286	833			

Table 38 ANOVA details for the undiversified and diversified samples. The suffixes *, ** and *** indicate that differences between the high and low coherence groups are significant at the 0.10, .05 and .01 significance level respectively.

Test of Homogeneity of Variances

		Levene Statistic	**df1**	**df2**	**Sig.**
Undiversified	**Sales growth**	8.447	1	3226	.004*
	ROA	.165	1	3299	.685
	EBIT margin	23.914	1	3267	.000*
	P/E ratio	3.953	1	1989	.047*
Diversified	**Sales growth**	5.156	1	1212	.023*
	ROA	10.251	1	1248	.001*
	EBIT margin	2.773	1	1231	.096
	P/E ratio	2.835	1	832	.093

Table 39 Levene's test of homogeneity of variances for the three size segment samples. Suffix * indicates that the variances between the two coherence groups are unequal at the 0.05 significance level. This means that the ANOVA assumption of equal variances is not met and this can influence the ANOVA results.

Independent Samples Kruskal-Wallis Test (Non-parametric One-way ANOVA)		
numSegs (Binned)		**Sig.**
Undiversified	**Sales growth**	.000***
	ROA	.000***
	EBIT margin	.000***
	P/E ratio	.000***
Diversified	**Sales growth**	.344
	ROA	.000***
	EBIT margin	.000***
	P/E ratio	.276

Table 40 Kruskal-Wallis Test (Non-parametric One-way ANOVA) for the full large sample. The suffixes *, ** and *** indicate that differences between the high and low coherence groups are significant at the 0.10, .05 and .01 significance level respectively.

3. Industry

The full ANOVA and Kruskal-Wallis results for the 2-digit SIC codes are given here.

Descriptives

SICH-2			N	Mean	Std. Deviation	Std. Error	95% Confidence Interval for Mean		Min.	Max.
							Lower Bound	Upper Bound		
20	**Sales growth**	Low Coherence	252	.09095	.306583	.019313	.05292	.12899	-.705	2.540
		High Coherence	338	.08136	.245759	.013368	.05507	.10765	-1.000	2.039
		Total	590	.08546	.273192	.011247	.06337	.10755	-1.000	2.540
	ROA	Low Coherence	258	-.00013	.207161	.012897	-.02553	.02527	-.868	2.372
		High Coherence	348	.03014	.136189	.007301	.01578	.04450	-1.184	.530
		Total	606	.01725	.170566	.006929	.00364	.03086	-1.184	2.372
	EBIT margin	Low Coherence	259	.04668	.085846	.005334	.03618	.05719	-.296	.394
		High Coherence	341	.06272	.119835	.006489	.04996	.07549	-1.004	.366
		Total	600	.05580	.106717	.004357	.04724	.06436	-1.004	.394
	P/E ratio	Low Coherence	130	25.948	33.7512	2.960174	20.091	31.804	1.161	310.10
		High Coherence	244	25.945	29.7503	1.904567	22.194	29.697	1.218	307.96
		Total	374	25.946	31.1540	1.610936	22.778	29.114	1.161	310.10
22	**Sales growth**	Low Coherence	76	-.01272	.204624	.023472	-.05948	.03404	-.525	.925
		High Coherence	103	.00666	.218270	.021507	-.03600	.04932	-.661	.669
		Total	179	-.00157	.212214	.015862	-.03287	.02973	-.661	.925
	ROA	Low Coherence	78	-.06676	.338542	.038332	-.14309	.00957	-1.490	.468
		High Coherence	105	-.00273	.119211	.011634	-.02580	.02034	-.763	.246
		Total	183	-.03002	.240037	.017744	-.06503	.00499	-1.490	.468
	EBIT margin	Low Coherence	78	.04598	.092652	.010491	.02509	.06687	-.318	.227
		High Coherence	105	.04110	.155722	.015197	.01097	.07124	-1.303	.203
		Total	183	.04318	.132267	.009777	.02389	.06247	-1.303	.227
	P/E ratio	Low Coherence	33	13.490	14.2069	2.473104	8.452	18.527	1.300	65.33
		High Coherence	58	21.197	24.6180	3.232496	14.724	27.670	3.374	122.82
		Total	91	18.402	21.6673	2.271354	13.889	22.914	1.300	122.82
23	**Sales growth**	Low Coherence	173	.06918	.363029	.027601	.01470	.12366	-.578	3.365
		High Coherence	198	.09937	.327196	.023253	.05351	.14522	-.708	3.341
		Total	371	.08529	.344228	.017871	.05015	.12043	-.708	3.365
	ROA	Low Coherence	174	-.03926	.354426	.026869	-.09230	.01377	-1.955	2.170
		High Coherence	200	.02814	.157828	.011160	.00613	.05015	-.827	.383
		Total	374	-.00322	.269602	.013941	-.03063	.02419	-1.955	2.170
	EBIT margin	Low Coherence	175	.03982	.090926	.006873	.02625	.05339	-.441	.355
		High Coherence	198	.07596	.077526	.005510	.06510	.08683	-.263	.295
		Total	373	.05901	.085885	.004447	.05026	.06775	-.441	.355

	P/E ratio	Low Coherence	93	23.822	37.3637	3.874438	16.127	31.517	3.228	320.28
		High Coherence	141	18.508	16.3950	1.380712	15.778	21.238	1.222	107.74
		Total	234	20.620	26.8240	1.753542	17.165	24.075	1.222	320.28
25	**Sales growth**	Low Coherence	7	.00060	.158091	.059753	-.14561	.14681	-.272	.217
		High Coherence	91	.06527	.197712	.020726	.02410	.10645	-.343	.831
		Total	98	.06065	.195180	.019716	.02152	.09978	-.343	.831
	ROA	Low Coherence	7	-.16603	.485141	.183366	-.61471	.28265	-1.243	.121
		High Coherence	91	.04010	.109242	.011452	.01735	.06285	-.620	.194
		Total	98	.02538	.168755	.017047	-.00846	.05921	-1.243	.194
	EBIT margin	Low Coherence	7	.05983	.081146	.030670	-.01522	.13487	-.074	.174
		High Coherence	91	.07598	.057050	.005980	.06410	.08786	-.140	.175
		Total	98	.07483	.058691	.005929	.06306	.08660	-.140	.175
	P/E ratio	Low Coherence	5	41.651	70.8951	31.705271	-46.377	129.679	6.699	168.29
		High Coherence	66	22.313	22.3961	2.756774	16.808	27.819	1.464	167.06
		Total	71	23.675	27.8890	3.309815	17.074	30.276	1.464	168.29
26	**Sales growth**	Low Coherence	77	.12018	.297772	.033934	.05259	.18777	-.355	1.251
		High Coherence	176	.04453	.181617	.013690	.01751	.07155	-.623	.886
		Total	253	.06755	.225530	.014179	.03963	.09548	-.623	1.251
	ROA	Low Coherence	82	-.01327	.136227	.015044	-.04320	.01666	-.613	.307
		High Coherence	181	.01429	.145998	.010852	-.00712	.03571	-1.296	.252
		Total	263	.00570	.143336	.008838	-.01170	.02310	-1.296	.307
	EBIT margin	Low Coherence	82	.05725	.082411	.009101	.03915	.07536	-.264	.257
		High Coherence	179	.07305	.102161	.007636	.05798	.08812	-.485	.249
		Total	261	.06809	.096515	.005974	.05632	.07985	-.485	.257
	P/E ratio	Low Coherence	37	21.776	23.6614	3.889902	13.887	29.665	1.289	118.18
		High Coherence	127	20.012	23.1467	2.053933	15.947	24.077	1.696	241.54
		Total	164	20.410	23.2023	1.811798	16.832	23.988	1.289	241.54
27	**Sales growth**	Low Coherence	4	-.05326	.081504	.040752	-.18296	.07643	-.137	.045
		High Coherence	37	.04333	.184713	.030367	-.01825	.10492	-.325	.496
		Total	41	.03391	.179018	.027958	-.02259	.09042	-.325	.496
	ROA	Low Coherence	4	.00307	.079769	.039884	-.12386	.13000	-.105	.087
		High Coherence	37	.02492	.080515	.013237	-.00192	.05177	-.268	.132
		Total	41	.02279	.079717	.012450	-.00237	.04795	-.268	.132
	EBIT margin	Low Coherence	4	.02702	.077985	.038993	-.09707	.15111	-.065	.125
		High Coherence	37	.06632	.063331	.010412	.04521	.08744	-.096	.158
		Total	41	.06249	.064848	.010128	.04202	.08296	-.096	.158
	P/E ratio	Low Coherence	3	15.234	2.4806	1.432194	9.072	21.396	12.500	17.34
		High Coherence	26	27.885	61.0775	11.978280	3.215	52.555	7.472	325.85
		Total	29	26.576	57.8496	10.742407	4.571	48.581	7.472	325.85
30	**Sales growth**	Low Coherence	15	.09179	.343599	.088717	-.09849	.28207	-.232	1.166
		High Coherence	24	.12963	.335452	.068474	-.01202	.27128	-.305	1.389
		Total	39	.11508	.334593	.053578	.00662	.22354	-.305	1.389
	ROA	Low Coherence	16	.01921	.117711	.029428	-.04351	.08194	-.286	.218
		High Coherence	25	.04682	.122600	.024520	-.00379	.09742	-.406	.268
		Total	41	.03604	.120001	.018741	-.00183	.07392	-.406	.268
	EBIT margin	Low Coherence	16	.06532	.066494	.016623	.02988	.10075	-.032	.248
		High Coherence	25	.07812	.071640	.014328	.04855	.10769	-.070	.279
		Total	41	.07312	.069119	.010795	.05131	.09494	-.070	.279
	P/E ratio	Low Coherence	9	15.735	6.1966	2.065528	10.972	20.498	5.473	28.31
		High Coherence	21	18.779	8.9097	1.944266	14.724	22.835	8.251	41.71
		Total	30	17.866	8.2069	1.498373	14.802	20.931	5.473	41.71
31	**Sales growth**	Low Coherence	50	.08666	.554896	.078474	-.07104	.24436	-.995	3.282
		High Coherence	97	.10152	.199829	.020290	.06125	.14180	-.238	.869
		Total	147	.09647	.360064	.029698	.03777	.15516	-.995	3.282
	ROA	Low Coherence	48	.04368	.181610	.026213	-.00905	.09642	-.717	.310
		High Coherence	97	.07928	.124372	.012628	.05421	.10434	-.621	.344

		Total	145	.06749	.146150	.012137	.04350	.09148	-.717	.344
	EBIT margin	Low Coherence	49	.07093	.121781	.017397	.03595	.10591	-.256	.362
		High Coherence	97	.09933	.081207	.008245	.08296	.11570	-.059	.380
		Total	146	.08980	.097245	.008048	.07389	.10571	-.256	.380
	P/E ratio	Low Coherence	31	15.896	9.8526	1.769588	12.282	19.510	2.111	36.09
		High Coherence	83	17.075	11.1655	1.225575	14.637	19.513	5.257	82.13
		Total	114	16.754	10.7943	1.010981	14.751	18.757	2.111	82.13
39	**Sales growth**	Low Coherence	6	.00699	.061515	.025113	-.05756	.07155	-.055	.124
		High Coherence	15	-.00405	.099813	.025772	-.05933	.05122	-.208	.116
		Total	21	-.00090	.089140	.019452	-.04147	.03968	-.208	.124
	ROA	Low Coherence	6	-.02531	.057219	.023360	-.08536	.03474	-.110	.023
		High Coherence	15	.02687	.024416	.006304	.01335	.04039	-.008	.067
		Total	21	.01196	.042652	.009307	-.00745	.03138	-.110	.067
	EBIT margin	Low Coherence	6	.03220	.047565	.019418	-.01772	.08211	-.041	.075
		High Coherence	15	.05358	.026488	.006839	.03891	.06825	.000	.095
		Total	21	.04747	.033981	.007415	.03200	.06294	-.041	.095
	P/E ratio	Low Coherence	3	27.912	25.1515	14.521225	-34.568	90.392	7.464	56.00
		High Coherence	13	38.646	42.0827	11.671641	13.216	64.077	7.568	157.15
		Total	16	36.634	38.9850	9.746257	15.860	57.407	7.464	157.15
52	**Sales growth**	Low Coherence	10	-.03503	.192332	.060821	-.17262	.10255	-.424	.164
		High Coherence	42	.07553	.220116	.033965	.00694	.14412	-.662	.854
		Total	52	.05427	.217749	.030196	-.00635	.11489	-.662	.854
	ROA	Low Coherence	11	.02881	.482253	.145405	-.29517	.35280	-.658	1.309
		High Coherence	43	.02523	.140171	.021376	-.01791	.06837	-.588	.214
		Total	54	.02596	.243830	.033181	-.04059	.09251	-.658	1.309
	EBIT margin	Low Coherence	11	.02308	.060551	.018257	-.01759	.06376	-.047	.154
		High Coherence	43	.04630	.046075	.007026	.03212	.06047	-.079	.196
		Total	54	.04157	.049630	.006754	.02802	.05511	-.079	.196
	P/E ratio	Low Coherence	4	17.049	12.1832	6.091578	-2.337	36.436	3.955	33.30
		High Coherence	29	23.433	15.2382	2.829669	17.637	29.229	5.826	66.25
		Total	33	22.659	14.8852	2.591186	17.381	27.937	3.955	66.25
53	**Sales growth**	Low Coherence	64	.07676	.278995	.034874	.00707	.14645	-.544	1.588
		High Coherence	107	.07219	.146099	.014124	.04419	.10019	-.393	.751
		Total	171	.07390	.205329	.015702	.04291	.10490	-.544	1.588
	ROA	Low Coherence	64	-.01515	.146321	.018290	-.05170	.02140	-.688	.198
		High Coherence	107	.02081	.233270	.022551	-.02389	.06552	-2.265	.186
		Total	171	.00736	.205349	.015703	-.02364	.03835	-2.265	.198
	EBIT margin	Low Coherence	64	.02025	.073174	.009147	.00197	.03852	-.365	.128
		High Coherence	107	.04788	.041043	.003968	.04001	.05575	-.168	.138
		Total	171	.03754	.056697	.004336	.02898	.04610	-.365	.138
	P/E ratio	Low Coherence	43	31.755	39.1482	5.970044	19.707	43.803	5.302	239.51
		High Coherence	89	18.276	10.5754	1.120995	16.049	20.504	3.085	65.58
		Total	132	22.667	24.6312	2.143874	18.426	26.908	3.085	239.51
54	**Sales growth**	Low Coherence	101	.00976	.190821	.018987	-.02791	.04743	-1.000	1.268
		High Coherence	95	.10689	.238731	.024493	.05826	.15552	-.164	1.512
		Total	196	.05684	.220261	.015733	.02581	.08787	-1.000	1.512
	ROA	Low Coherence	102	-.01585	.133476	.013216	-.04207	.01037	-.540	.401
		High Coherence	97	-.00400	.202789	.020590	-.04487	.03687	-1.620	.137
		Total	199	-.01007	.170475	.012085	-.03390	.01376	-1.620	.401
	EBIT margin	Low Coherence	101	.02077	.026845	.002671	.01547	.02607	-.039	.103
		High Coherence	96	.03064	.107550	.010977	.00885	.05243	-.887	.417
		Total	197	.02558	.077451	.005518	.01470	.03646	-.887	.417
	P/E ratio	Low Coherence	38	20.712	15.5841	2.528070	15.590	25.835	2.525	67.62
		High Coherence	66	34.234	57.1985	7.040648	20.173	48.295	2.950	374.42
		Total	104	29.293	46.8475	4.593780	20.183	38.404	2.525	374.42
55	**Sales growth**	Low Coherence	133	.15638	.324111	.028104	.10079	.21198	-.422	2.069
		High Coherence	41	.37711	.662260	.103428	.16807	.58614	-.405	3.357
		Total	174	.20839	.436332	.033078	.14311	.27368	-.422	3.357

	ROA	Low Coherence	134	.00222	.115246	.009956	-.01747	.02191	-1.063	.125
		High Coherence	43	.00452	.139318	.021246	-.03836	.04740	-.776	.141
		Total	177	.00278	.121118	.009104	-.01519	.02074	-1.063	.141
	EBIT margin	Low Coherence	133	.03133	.043192	.003745	.02392	.03874	-.175	.243
		High Coherence	42	.04439	.060009	.009260	.02569	.06309	-.121	.179
		Total	175	.03447	.047907	.003621	.02732	.04161	-.175	.243
	P/E ratio	Low Coherence	90	21.920	56.6021	5.966381	10.065	33.775	1.597	522.43
		High Coherence	32	23.508	26.3562	4.659163	14.006	33.011	2.919	152.76
		Total	122	22.336	50.3485	4.558337	13.312	31.361	1.597	522.43
56	**Sales growth**	Low Coherence	263	.10160	.179612	.011075	.07979	.12340	-.390	1.136
		High Coherence	125	.09686	.201538	.018026	.06118	.13254	-1.000	.846
		Total	388	.10007	.186708	.009479	.08143	.11871	-1.000	1.136
	ROA	Low Coherence	268	.03197	.167785	.010249	.01179	.05215	-1.214	.290
		High Coherence	125	.05242	.129717	.011602	.02946	.07539	-.543	.327
		Total	393	.03848	.156807	.007910	.02293	.05403	-1.214	.327
	EBIT margin	Low Coherence	268	.05552	.062589	.003823	.04799	.06305	-.197	.222
		High Coherence	124	.06398	.062428	.005606	.05289	.07508	-.126	.232
		Total	392	.05820	.062582	.003161	.05198	.06441	-.197	.232
	P/E ratio	Low Coherence	201	20.920	18.3974	1.297651	18.361	23.479	1.909	199.43
		High Coherence	98	22.721	48.1229	4.861146	13.073	32.369	6.817	485.88
		Total	299	21.510	31.3318	1.811962	17.945	25.076	1.909	485.88
57	**Sales growth**	Low Coherence	78	.09765	.333041	.037709	.02256	.17274	-.355	2.119
		High Coherence	82	.07237	.206059	.022755	.02709	.11765	-.220	1.139
		Total	160	.08469	.274783	.021723	.04179	.12760	-.355	2.119
	ROA	Low Coherence	78	-.04462	.206689	.023403	-.09122	.00199	-1.042	.252
		High Coherence	80	.00012	.178636	.019972	-.03963	.03987	-.708	1.091
		Total	158	-.02196	.193681	.015408	-.05240	.00847	-1.042	1.091
	EBIT margin	Low Coherence	78	.01118	.054739	.006198	-.00116	.02352	-.181	.125
		High Coherence	82	.02111	.057910	.006395	.00839	.03383	-.159	.154
		Total	160	.01627	.056429	.004461	.00746	.02508	-.181	.154
	P/E ratio	Low Coherence	40	17.007	17.0773	2.700162	11.545	22.469	2.128	98.46
		High Coherence	44	27.706	47.9727	7.232159	13.121	42.291	4.939	327.65
		Total	84	22.611	36.8539	4.021091	14.613	30.609	2.128	327.65
58	**Sales growth**	Low Coherence	415	.08481	.242902	.011924	.06137	.10825	-.819	2.204
		High Coherence	213	.13002	.272361	.018662	.09324	.16681	-.592	2.021
		Total	628	.10014	.253966	.010134	.08024	.12005	-.819	2.204
	ROA	Low Coherence	418	.00179	.178194	.008716	-.01534	.01892	-1.379	1.655
		High Coherence	220	-.00161	.168770	.011378	-.02403	.02082	-.941	.443
		Total	638	.00062	.174876	.006923	-.01298	.01421	-1.379	1.655
	EBIT margin	Low Coherence	423	.04288	.066844	.003250	.03649	.04927	-.374	.235
		High Coherence	216	.05939	.097986	.006667	.04625	.07254	-.311	.288
		Total	639	.04846	.079070	.003128	.04232	.05460	-.374	.288
	P/E ratio	Low Coherence	220	20.768	28.2569	1.905082	17.014	24.523	1.468	358.02
		High Coherence	148	29.034	33.4976	2.753486	23.592	34.475	6.531	288.78
		Total	368	24.093	30.6982	1.600254	20.946	27.239	1.468	358.02
59	**Sales growth**	Low Coherence	401	.14943	.307551	.015358	.11923	.17962	-1.000	1.846
		High Coherence	373	.23875	.509145	.026363	.18691	.29058	-.958	3.249
		Total	774	.19247	.419156	.015066	.16289	.22205	-1.000	3.249
	ROA	Low Coherence	409	-.01864	.286443	.014164	-.04648	.00920	-3.026	.525
		High Coherence	403	-.08652	.340836	.016978	-.11990	-.05314	-2.630	.570
		Total	812	-.05233	.316250	.011098	-.07411	-.03054	-3.026	.570
	EBIT margin	Low Coherence	400	.04503	.101004	.005050	.03510	.05496	-.850	.233
		High Coherence	381	-.01687	.186160	.009537	-.03562	.00188	-1.278	.210
		Total	781	.01483	.151857	.005434	.00417	.02550	-1.278	.233
	P/E ratio	Low Coherence	246	27.817	48.0075	3.060844	21.788	33.846	2.081	530.13
		High Coherence	217	33.039	38.5586	2.617528	27.879	38.198	3.009	300.60
		Total	463	30.264	43.8647	2.038566	26.258	34.270	2.081	530.13
70	**Sales**	Low Coherence	96	.05728	.502690	.051306	-.04457	.15914	-1.000	3.165

	growth	High Coherence	64	.32841	.708996	.088625	.15131	.50551	-.415	2.852
		Total	160	.16573	.606555	.047952	.07103	.26044	-1.000	3.165
	ROA	Low Coherence	106	-.02260	.221190	.021484	-.06520	.01999	-1.763	.335
		High Coherence	71	-.00025	.119066	.014131	-.02843	.02793	-.939	.083
		Total	177	-.01364	.186942	.014051	-.04137	.01409	-1.763	.335
	EBIT margin	Low Coherence	100	.08182	.142112	.014211	.05362	.11002	-.429	.460
		High Coherence	69	.13606	.145229	.017484	.10117	.17094	-.421	.692
		Total	169	.10396	.145441	.011188	.08188	.12605	-.429	.692
	P/E ratio	Low Coherence	45	23.570	23.6414	3.524256	16.467	30.673	1.387	104.42
		High Coherence	52	40.569	79.4794	11.021807	18.442	62.696	3.759	451.07
		Total	97	32.683	60.7015	6.163305	20.449	44.917	1.387	451.07

Table 41 Descriptives of the two-digit SIC code industry samples.

ANOVA

			Sum of Squares	df	Mean Square	F	Sig.
30	**Sales growth**	Between Groups	.013	1	.013	.115	.736
		Within Groups	4.241	37	.115		
		Total	4.254	38			
	ROA	Between Groups	.007	1	.007	.510	.479
		Within Groups	.569	39	.015		
		Total	.576	40			
	EBIT margin	Between Groups	.002	1	.002	.329	.569
		Within Groups	.189	39	.005		
		Total	.191	40			
	P/E ratio	Between Groups	58.404	1	58.404	.863	.361
		Within Groups	1894.852	28	67.673		
		Total	1953.256	29			
31	**Sales growth**	Between Groups	.007	1	.007	.056	.813
		Within Groups	18.921	145	.130		
		Total	18.928	146			
	ROA	Between Groups	.041	1	.041	1.917	.168
		Within Groups	3.035	143	.021		
		Total	3.076	144			
	EBIT margin	Between Groups	.026	1	.026	2.811	.096*
		Within Groups	1.345	144	.009		
		Total	1.371	145			
	P/E ratio	Between Groups	31.387	1	31.387	.268	.606
		Within Groups	13135.091	112	117.278		
		Total	13166.478	113			
39	**Sales growth**	Between Groups	.001	1	.001	.063	.805
		Within Groups	.158	19	.008		
		Total	.159	20			
	ROA	Between Groups	.012	1	.012	8.969	.007***
		Within Groups	.025	19	.001		
		Total	.036	20			
	EBIT margin	Between Groups	.002	1	.002	1.762	.200
		Within Groups	.021	19	.001		
		Total	.023	20			
	P/E ratio	Between Groups	280.847	1	280.847	.175	.682
		Within Groups	22516.640	14	1608.331		
		Total	22797.487	15			
52	**Sales growth**	Between Groups	.099	1	.099	2.128	.151
		Within Groups	2.319	50	.046		
		Total	2.418	51			
	ROA	Between Groups	.000	1	.000	.002	.966
		Within Groups	3.151	52	.061		
		Total	3.151	53			
	EBIT	Between Groups	.005	1	.005	1.950	.169

	margin	Within Groups	.126	52	.002		
		Total	.131	53			
	P/E ratio	Between Groups	143.248	1	143.248	.639	.430
		Within Groups	6946.993	31	224.097		
		Total	7090.241	32			
53	**Sales growth**	Between Groups	.001	1	.001	.020	.889
		Within Groups	7.166	169	.042		
		Total	7.167	170			
	ROA	Between Groups	.052	1	.052	1.230	.269
		Within Groups	7.117	169	.042		
		Total	7.169	170			
	EBIT margin	Between Groups	.031	1	.031	10.019	.002***
		Within Groups	.516	169	.003		
		Total	.546	170			
	P/E ratio	Between Groups	5267.074	1	5267.074	9.227	.003***
		Within Groups	74210.341	130	570.849		
		Total	79477.415	131			
54	**Sales growth**	Between Groups	.462	1	.462	9.956	.002***
		Within Groups	8.999	194	.046		
		Total	9.460	195			
	ROA	Between Groups	.007	1	.007	.240	.625
		Within Groups	5.747	197	.029		
		Total	5.754	198			
	EBIT margin	Between Groups	.005	1	.005	.799	.372
		Within Groups	1.171	195	.006		
		Total	1.176	196			
	P/E ratio	Between Groups	4408.966	1	4408.966	2.029	.157
		Within Groups	221644.367	102	2172.984		
		Total	226053.333	103			
55	**Sales growth**	Between Groups	1.527	1	1.527	8.361	.004***
		Within Groups	31.410	172	.183		
		Total	32.937	173			
	ROA	Between Groups	.000	1	.000	.012	.914
		Within Groups	2.582	175	.015		
		Total	2.582	176			
	EBIT margin	Between Groups	.005	1	.005	2.388	.124
		Within Groups	.394	173	.002		
		Total	.399	174			
	P/E ratio	Between Groups	59.570	1	59.570	.023	.879
		Within Groups	306671.731	120	2555.598		
		Total	306731.302	121			
56	**Sales growth**	Between Groups	.002	1	.002	.054	.816
		Within Groups	13.489	386	.035		
		Total	13.491	387			
	ROA	Between Groups	.036	1	.036	1.452	.229
		Within Groups	9.603	391	.025		
		Total	9.639	392			
	EBIT margin	Between Groups	.006	1	.006	1.552	.214
		Within Groups	1.525	390	.004		
		Total	1.531	391			
	P/E ratio	Between Groups	213.782	1	213.782	.217	.642
		Within Groups	292326.459	297	984.264		
		Total	292540.241	298			
57	**Sales growth**	Between Groups	.026	1	.026	.337	.562
		Within Groups	11.980	158	.076		
		Total	12.005	159			
	ROA	Between Groups	.079	1	.079	2.122	.147
		Within Groups	5.810	156	.037		

		Total	5.889	157			
	EBIT margin	Between Groups	.004	1	.004	1.239	.267
		Within Groups	.502	158	.003		
		Total	.506	159			
	P/E ratio	Between Groups	2398.303	1	2398.303	1.782	.186
		Within Groups	110333.176	82	1345.527		
		Total	112731.479	83			
58	**Sales growth**	Between Groups	.288	1	.288	4.486	.035**
		Within Groups	40.153	626	.064		
		Total	40.441	627			
	ROA	Between Groups	.002	1	.002	.054	.816
		Within Groups	19.479	636	.031		
		Total	19.481	637			
	EBIT margin	Between Groups	.039	1	.039	6.291	.012**
		Within Groups	3.950	637	.006		
		Total	3.989	638			
	P/E ratio	Between Groups	6044.597	1	6044.597	6.510	.011**
		Within Groups	339808.513	366	928.439		
		Total	345853.110	367			
59	**Sales growth**	Between Groups	1.542	1	1.542	8.865	.003***
		Within Groups	134.268	772	.174		
		Total	135.810	773			
	ROA	Between Groups	.935	1	.935	9.449	.002***
		Within Groups	80.176	810	.099		
		Total	81.112	811			
	EBIT margin	Between Groups	.748	1	.748	33.783	.000***
		Within Groups	17.240	779	.022		
		Total	17.987	780			
	P/E ratio	Between Groups	3143.586	1	3143.586	1.636	.202
		Within Groups	885796.611	461	1921.468		
		Total	888940.197	462			
70	**Sales growth**	Between Groups	2.823	1	2.823	8.011	.005***
		Within Groups	55.675	158	.352		
		Total	58.497	159			
	ROA	Between Groups	.021	1	.021	.607	.437
		Within Groups	6.129	175	.035		
		Total	6.151	176			
	EBIT margin	Between Groups	.120	1	.120	5.841	.017**
		Within Groups	3.434	167	.021		
		Total	3.554	168			
	P/E ratio	Between Groups	6970.777	1	6970.777	1.910	.170
		Within Groups	346757.917	95	3650.083		
		Total	353728.694	96			

Table 42 ANOVA details for the two-digit SIC code industry samples. The suffixes *, ** and *** indicate that differences between the high and low coherence groups are significant at the 0.10, .05 and .01 significance level respectively.

Test of Homogeneity of Variances

	SICH-2	Levene Statistic	df1	df2	Sig.
20	**Sales growth**	2.130	1	588	.145
	ROA	2.415	1	604	.121
	EBIT margin	2.019	1	598	.156
	P/E ratio	.490	1	372	.485
22	**Sales growth**	1.214	1	177	.272
	ROA	18.243	1	181	.000*
	EBIT margin	.125	1	181	.724
	P/E ratio	1.458	1	89	.230

23	**Sales growth**	.253	1	369	.615
	ROA	12.195	1	372	.001*
	EBIT margin	.861	1	371	.354
	P/E ratio	5.883	1	232	.016*
25	**Sales growth**	.109	1	96	.742
	ROA	27.970	1	96	.000*
	EBIT margin	1.304	1	96	.256
	P/E ratio	15.785	1	69	.000*
26	**Sales growth**	13.839	1	251	.000*
	ROA	1.141	1	261	.286
	EBIT margin	.124	1	259	.725
	P/E ratio	1.996	1	162	.160
27	**Sales growth**	1.287	1	39	.263
	ROA	.052	1	39	.820
	EBIT margin	.039	1	39	.845
	P/E ratio	.440	1	27	.513
30	**Sales growth**	.135	1	37	.715
	ROA	.243	1	39	.625
	EBIT margin	.002	1	39	.969
	P/E ratio	1.055	1	28	.313
31	**Sales growth**	7.480	1	145	.007*
	ROA	6.920	1	143	.009*
	EBIT margin	6.298	1	144	.013*
	P/E ratio	.916	1	112	.341
39	**Sales growth**	3.844	1	19	.065
	ROA	13.492	1	19	.002*
	EBIT margin	5.671	1	19	.028*
	P/E ratio	.301	1	14	.592
52	**Sales growth**	.043	1	50	.837
	ROA	8.558	1	52	.005*
	EBIT margin	1.410	1	52	.240
	P/E ratio	.365	1	31	.550
53	**Sales growth**	7.450	1	169	.007*
	ROA	.580	1	169	.448
	EBIT margin	4.196	1	169	.042
	P/E ratio	16.511	1	130	.000*
54	**Sales growth**	1.514	1	194	.220
	ROA	.125	1	197	.724
	EBIT margin	1.893	1	195	.170
	P/E ratio	3.570	1	102	.062
55	**Sales growth**	12.168	1	172	.001*
	ROA	.488	1	175	.486
	EBIT margin	5.628	1	173	.019*
	P/E ratio	.047	1	120	.829
56	**Sales growth**	.068	1	386	.794
	ROA	1.241	1	391	.266
	EBIT margin	.146	1	390	.702
	P/E ratio	.756	1	297	.385
57	**Sales growth**	2.606	1	158	.108
	ROA	1.927	1	156	.167
	EBIT margin	.015	1	158	.903
	P/E ratio	.997	1	82	.321
58	**Sales growth**	.910	1	626	.340
	ROA	.927	1	636	.336
	EBIT margin	18.368	1	637	.000*
	P/E ratio	2.999	1	366	.084
59	**Sales growth**	19.724	1	772	.000*
	ROA	10.822	1	810	.001*

	EBIT margin	23.752	1	779	.000*
	P/E ratio	.768	1	461	.381
70	**Sales growth**	9.194	1	158	.003*
	ROA	3.522	1	175	.062
	EBIT margin	.209	1	167	.648
	P/E ratio	2.467	1	95	.120

Table 43 Levene's test of homogeneity of variances for the two-digit SIC code industry samples. Suffix * indicates that the variances between the two coherence groups are unequal at the 0.05 significance level. This means that the ANOVA assumption of equal variances is not met and this can influence the ANOVA results.

Independent Samples Kruskal-Wallis Test (Non-parametric One-way ANOVA)		
	SICH-2	Sig.
20	**Sales growth**	.074*
	ROA	.000***
	EBIT margin	.000***
	P/E ratio	.258
22	**Sales growth**	.336
	ROA	.793
	EBIT margin	.240
	P/E ratio	.006***
23	**Sales growth**	.026**
	ROA	.000***
	EBIT margin	.000***
	P/E ratio	.541
25	**Sales growth**	.404
	ROA	.233
	EBIT margin	.586
	P/E ratio	.271
26	**Sales growth**	.101
	ROA	.003***
	EBIT margin	.014
	P/E ratio	.447
27	**Sales growth**	.272
	ROA	.455
	EBIT margin	.356
	P/E ratio	1.000
30	**Sales growth**	.403
	ROA	.240
	EBIT margin	.539
	P/E ratio	.455
31	**Sales growth**	.237
	ROA	.164
	EBIT margin	.097*
	P/E ratio	.490
39	**Sales growth**	.938
	ROA	.043**
	EBIT margin	.350
	P/E ratio	.638
52	**Sales growth**	.104
	ROA	.233
	EBIT margin	.141
	P/E ratio	.408
53	**Sales growth**	.446
	ROA	.001***
	EBIT margin	.000***
	P/E ratio	.018**
54	**Sales growth**	.000***

	ROA	.000***
	EBIT margin	.000***
	P/E ratio	.075*
55	**Sales growth**	.017**
	ROA	.381
	EBIT margin	.034**
	P/E ratio	.054*
56	**Sales growth**	.457
	ROA	.194
	EBIT margin	.102
	P/E ratio	.525
57	**Sales growth**	.723
	ROA	.679
	EBIT margin	.266
	P/E ratio	.024**
58	**Sales growth**	.000***
	ROA	.013**
	EBIT margin	.000***
	P/E ratio	.000***
59	**Sales growth**	.257
	ROA	.000***
	EBIT margin	.000***
	P/E ratio	.004***
70	**Sales growth**	.003***
	ROA	.292
	EBIT margin	.014**
	P/E ratio	.056*

Table 44 Kruskal-Wallis Test (Non-parametric One-way ANOVA) for the two-digit SIC code industry samples. The suffixes *, ** and *** indicate that differences between the high and low coherence groups are significant at the 0.10, .05 and .01 significance level respectively.

C. Corrected for Relevant Factors Together

The full ANOVA and Kruskal-Wallis results for the three SIC-codes 20, 58 and 59, with the factors size and diversification included are given here.

Descriptives

SIC	Size (sales)	Diversification	Premium	Coherence Group	N	Mean	Std. Deviation	Std. Error	95% Confidence Interval for Mean		Min.	Max.
									Lower Bound	Upper Bound		
20	**Small**	**Undiversified**	**Sales growth**	Low	94	.04434	.251586	.025949	-.00719	.09587	-.705	.980
				High	105	.09046	.295822	.028869	.03321	.14771	-1.000	.762
				Total	199	.06868	.276094	.019572	.03008	.10727	-1.000	.980
			ROA	Low	96	-.03832	.316153	.032267	-.10238	.02573	-.868	2.372
				High	111	-.01682	.213551	.020269	-.05699	.02335	-1.184	.530
				Total	207	-.02679	.265635	.018463	-.06320	.00961	-1.184	2.372
			EBIT margin	Low	97	.01391	.103625	.010522	-.00698	.03479	-.296	.333
				High	106	.01295	.178245	.017313	-.02138	.04727	-1.004	.366
				Total	203	.01341	.147031	.010320	-.00694	.03375	-1.004	.366
			P/E ratio	Low	42	23.728	21.2355	3.276705	17.110	30.345	2.172	84.65
				High	58	28.243	42.1628	5.536246	17.156	39.329	1.798	307.96
				Total	100	26.346	34.8611	3.486109	19.429	33.264	1.798	307.96
		Diversified	**Sales growth**	Low	22	.19198	.553595	.118027	-.05347	.43743	-.146	2.540
				High	41	.06460	.254986	.039822	-.01588	.14508	-.873	.959
				Total	63	.10908	.386649	.048713	.01171	.20646	-.873	2.540
			ROA	Low	22	-.00196	.088125	.018788	-.04104	.03711	-.203	.213

				High	43	.01855	.073913	.011272	-.00420	.04129	-.197	.189
				Total	65	.01160	.078924	.009789	-.00795	.03116	-.203	.213
			EBIT margin	Low	22	.05423	.077061	.016429	.02006	.08840	-.067	.230
				High	41	.03664	.083448	.013032	.01030	.06298	-.116	.234
				Total	63	.04278	.081089	.010216	.02236	.06320	-.116	.234
			P/E ratio	Low	11	33.576	44.6701	13.46854	3.566	63.585	7.042	159.50
				High	26	23.222	17.7918	3.489261	16.035	30.408	3.817	71.63
				Total	37	26.300	28.2334	4.641547	16.886	35.713	3.817	159.50
	Medium	**Undiversified**	**Sales growth**	Low	50	.05925	.166755	.023583	.01186	.10664	-.456	.580
				High	68	.11631	.252984	.030679	.05508	.17755	-.058	2.039
				Total	118	.09213	.221580	.020398	.05173	.13253	-.456	2.039
			ROA	Low	52	.01742	.129964	.018023	-.01876	.05361	-.292	.725
				High	69	.04604	.044572	.005366	.03534	.05675	-.101	.137
				Total	121	.03374	.092231	.008385	.01714	.05034	-.292	.725
			EBIT margin	Low	52	.08666	.080093	.011107	.06436	.10896	-.066	.394
				High	69	.05860	.036876	.004439	.04974	.06746	-.044	.172
				Total	121	.07066	.060758	.005523	.05972	.08160	-.066	.394
			P/E ratio	Low	29	25.629	23.6518	4.392027	16.632	34.625	1.885	125.28
				High	55	31.176	35.1095	4.734159	21.684	40.667	3.865	210.27
				Total	84	29.261	31.5870	3.446420	22.406	36.116	1.885	210.27
		Diversified	**Sales growth**	Low	6	.02303	.077491	.031636	-.05830	.10435	-.066	.119
				High	18	.03168	.222322	.052402	-.07888	.14224	-.384	.630
				Total	24	.02952	.194559	.039714	-.05264	.11167	-.384	.630
			ROA	Low	6	.00877	.049580	.020241	-.04326	.06080	-.070	.076
				High	18	.01075	.087098	.020529	-.03257	.05406	-.244	.100
				Total	24	.01025	.078372	.015998	-.02284	.04334	-.244	.100
			EBIT margin	Low	6	.08563	.090465	.036932	-.00930	.18057	-.068	.178
				High	18	.07584	.051835	.012218	.05006	.10162	-.046	.166
				Total	24	.07829	.061513	.012556	.05231	.10426	-.068	.178
			P/E ratio	Low	3	30.080	9.2942	5.366024	6.992	53.168	20.054	38.41
				High	10	24.022	12.1601	3.845352	15.323	32.721	10.749	50.27
				Total	13	25.420	11.5046	3.190790	18.468	32.372	10.749	50.27
	Large	**Undiversified**	**Sales growth**	Low	40	.11410	.186523	.029492	.05445	.17376	-.104	.738
				High	58	.06427	.176347	.023156	.01790	.11064	-.365	1.048
				Total	98	.08461	.181296	.018314	.04826	.12096	-.365	1.048
			ROA	Low	41	.03656	.071243	.011126	.01408	.05905	-.167	.253
				High	59	.07672	.066775	.008693	.05932	.09412	-.095	.289
				Total	100	.06026	.071113	.007111	.04615	.07437	-.167	.289
			EBIT margin	Low	41	.06432	.051935	.008111	.04792	.08071	-.067	.204
				High	59	.13987	.074579	.009709	.12043	.15930	-.047	.281
				Total	100	.10889	.075783	.007578	.09385	.12393	-.067	.281
			P/E ratio	Low	28	20.524	15.4931	2.927911	14.516	26.531	1.161	81.58
				High	52	24.658	21.7939	3.022267	18.591	30.726	1.218	127.65
				Total	80	23.211	19.8142	2.215296	18.802	27.621	1.161	127.65
		Diversified	**Sales growth**	Low	40	.17159	.445268	.070403	.02918	.31399	-.618	2.446
				High	48	.06554	.183029	.026418	.01239	.11868	-.339	.850
				Total	88	.11374	.331353	.035322	.04353	.18395	-.618	2.446
			ROA	Low	41	.03004	.033510	.005233	.01946	.04062	-.051	.141
				High	48	.07626	.058800	.008487	.05919	.09333	-.048	.223
				Total	89	.05497	.053794	.005702	.04364	.06630	-.051	.223
			EBIT margin	Low	41	.04615	.038521	.006016	.03399	.05830	-.006	.197
				High	48	.10110	.043923	.006340	.08835	.11385	.012	.185
				Total	89	.07578	.049636	.005261	.06533	.08624	-.006	.197
			P/E ratio	Low	17	35.244	72.0222	17.46794	-1.786	72.275	5.918	310.10
				High	43	19.807	15.4036	2.349034	15.066	24.547	6.536	108.33
				Total	60	24.181	40.3090	5.203868	13.768	34.593	5.918	310.10
58	**Small**	**Undiver**	**Sales**	Low	177	.09326	.275460	.020705	.05240	.13412	-.819	2.204

	sified	growth	High	63	.19288	.414133	.052176	.08859	.29718	-.592	2.021
			Total	240	.11941	.319840	.020646	.07874	.16008	-.819	2.204
		ROA	Low	177	-.02075	.181067	.013610	-.04761	.00611	-1.379	.703
			High	64	-.06015	.177796	.022224	-.10456	-.01574	-.653	.170
			Total	241	-.03121	.180678	.011639	-.05414	-.00829	-1.379	.703
		EBIT	Low	182	.01812	.076989	.005707	.00686	.02938	-.374	.212
		margin	High	63	.00943	.088170	.011108	-.01277	.03164	-.282	.135
			Total	245	.01589	.079917	.005106	.00583	.02594	-.374	.212
		P/E	Low	86	18.592	15.5714	1.679111	15.253	21.930	1.468	83.68
		ratio	High	33	33.358	38.9439	6.779260	19.549	47.167	6.531	184.83
			Total	119	22.687	25.1001	2.300920	18.130	27.243	1.468	184.83
	Diversifi	Sales	Low	38	.05276	.235885	.038266	-.02477	.13029	-.551	1.001
	ed	growth	High	12	.04976	.290353	.083818	-.13473	.23424	-.482	.805
			Total	50	.05204	.246865	.034912	-.01812	.12220	-.551	1.001
		ROA	Low	38	.05403	.105902	.017180	.01922	.08884	-.177	.349
			High	17	-.08492	.204040	.049487	-.18983	.01998	-.556	.209
			Total	55	.01108	.155627	.020985	-.03099	.05315	-.556	.349
		EBIT	Low	38	.04311	.020964	.003401	.03622	.05000	.000	.092
		margin	High	14	-.03174	.142809	.038167	-.11419	.05072	-.311	.146
			Total	52	.02296	.081493	.011301	.00027	.04564	-.311	.146
		P/E	Low	23	9.718	6.4031	1.335138	6.949	12.487	1.575	27.88
		ratio	High	6	33.370	31.0582	12.67947	0.776	65.964	7.440	88.07
			Total	29	14.611	17.3073	3.213885	8.028	21.195	1.575	88.07
Mediu	Undiver	Sales	Low	131	.09713	.230317	.020123	.05732	.13694	-.292	1.270
m	sified	growth	High	75	.13618	.205330	.023709	.08894	.18342	-.362	1.106
			Total	206	.11135	.221839	.015456	.08088	.14182	-.362	1.270
		ROA	Low	133	.00363	.200486	.017384	-.03075	.03802	-1.024	1.655
			High	75	.01118	.191228	.022081	-.03281	.05518	-.941	.443
			Total	208	.00636	.196767	.013643	-.02054	.03325	-1.024	1.655
		EBIT	Low	133	.06288	.057494	.004985	.05302	.07274	-.202	.235
		margin	High	75	.07480	.083963	.009695	.05548	.09412	-.263	.246
			Total	208	.06718	.068271	.004734	.05785	.07651	-.263	.246
		P/E	Low	70	28.291	45.2559	5.409112	17.500	39.082	4.356	358.02
		ratio	High	56	33.603	42.8461	5.725547	22.128	45.077	6.818	288.78
			Total	126	30.652	44.1057	3.929251	22.875	38.428	4.356	358.02
	Diversifi	Sales	Low	20	.08439	.160201	.035822	.00942	.15937	-.079	.381
	ed	growth	High	8	.09679	.123106	.043525	-.00613	.19971	-.019	.316
			Total	28	.08794	.148397	.028044	.03039	.14548	-.079	.381
		ROA	Low	21	.00754	.086100	.018789	-.03165	.04673	-.166	.124
			High	8	-.02283	.122553	.043329	-.12529	.07962	-.182	.127
			Total	29	-.00084	.096130	.017851	-.03740	.03573	-.182	.127
		EBIT	Low	21	.06365	.039771	.008679	.04554	.08175	.020	.140
		margin	High	8	.03659	.056024	.019808	-.01025	.08342	-.061	.097
			Total	29	.05618	.045454	.008441	.03889	.07347	-.061	.140
		P/E	Low	9	25.561	15.3696	5.123215	13.747	37.375	6.977	46.89
		ratio	High	3	27.713	10.7353	6.198052	1.045	54.381	16.097	37.27
			Total	12	26.099	13.9177	4.017691	17.256	34.942	6.977	46.89
Large	Undiver	Sales	Low	42	.03700	.156384	.024131	-.01174	.08573	-.352	.646
	sified	growth	High	36	.07317	.091373	.015229	.04225	.10408	-.125	.319
			Total	78	.05369	.130944	.014826	.02417	.08321	-.352	.646
		ROA	Low	42	.03088	.178992	.027619	-.02490	.08666	-.915	.279
			High	36	.08711	.035922	.005987	.07496	.09926	-.022	.158
			Total	78	.05683	.135800	.015376	.02622	.08745	-.915	.279
		EBIT	Low	42	.07225	.051661	.007971	.05615	.08835	-.049	.207
		margin	High	36	.14997	.070373	.011729	.12616	.17378	.031	.288
			Total	78	.10812	.072062	.008159	.09188	.12437	-.049	.288

			P/E ratio	Low	29	17.151	8.8396	1.641469	13.789	20.513	4.810	46.75
				High	35	20.929	10.7531	1.817608	17.235	24.623	8.924	68.37
				Total	64	19.217	10.0362	1.254523	16.710	21.724	4.810	68.37
		Diversified	**Sales growth**	Low	7	.10249	.276019	.104326	-.15279	.35776	-.087	.709
				High	19	.06969	.082985	.019038	.02969	.10969	-.030	.253
				Total	26	.07852	.153177	.030040	.01665	.14039	-.087	.709
			ROA	Low	7	.06115	.074536	.028172	-.00779	.13008	-.010	.167
				High	20	.05736	.030559	.006833	.04306	.07166	-.004	.128
				Total	27	.05834	.044355	.008536	.04080	.07589	-.010	.167
			EBIT margin	Low	7	.06674	.028873	.010913	.04003	.09344	.035	.104
				High	20	.06888	.019896	.004449	.05957	.07819	.041	.110
				Total	27	.06832	.021968	.004228	.05963	.07701	.035	.110
			P/E ratio	Low	3	12.948	1.7336	1.000907	8.641	17.254	10.950	14.05
				High	15	19.906	8.4907	2.192286	15.204	24.608	10.748	38.03
				Total	18	18.746	8.1758	1.927054	14.681	22.812	10.748	38.03
59	**Small**	**Undiversified**	**Sales growth**	Low	120	.16283	.391401	.035730	.09208	.23358	-1.000	1.836
				High	112	.32364	.705199	.066635	.19160	.45568	-.958	3.249
				Total	232	.24046	.569533	.037392	.16679	.31413	-1.000	3.249
			ROA	Low	124	-.04504	.334489	.030038	-.10449	.01442	-2.470	.407
				High	132	-.21365	.464412	.040422	-.29361	-.13369	-2.630	.570
				Total	256	-.13198	.414602	.025913	-.18301	-.08095	-2.630	.570
			EBIT margin	Low	120	.00564	.148384	.013546	-.02118	.03246	-.850	.231
				High	113	-.11380	.289361	.027221	-.16774	-.05987	-1.278	.136
				Total	233	-.05229	.235146	.015405	-.08264	-.02193	-1.278	.231
			P/E ratio	Low	66	37.361	62.2327	7.660312	22.062	52.660	3.230	372.68
				High	42	27.032	21.2303	3.275914	20.416	33.648	3.009	84.28
				Total	108	33.344	50.5074	4.860081	23.709	42.979	3.009	372.68
		Diversified	**Sales growth**	Low	22	.05368	.387148	.082540	-.11797	.22534	-.524	1.110
				High	31	.27437	.409021	.073462	.12434	.42440	-.345	1.926
				Total	53	.18277	.411218	.056485	.06942	.29611	-.524	1.926
			ROA	Low	23	-.37663	.740765	.154460	-.69696	-.05630	-3.026	.068
				High	37	-.09226	.232858	.038282	-.16990	-.01463	-.905	.190
				Total	60	-.20127	.507087	.065465	-.33227	-.07028	-3.026	.190
			EBIT margin	Low	18	-.00972	.134376	.031673	-.07655	.05710	-.504	.090
				High	34	-.02848	.134376	.023045	-.07537	.01841	-.370	.171
				Total	52	-.02199	.133357	.018493	-.05911	.01514	-.504	.171
			P/E ratio	Low	5	15.766	6.1389	2.745378	8.144	23.388	7.568	23.14
				High	16	38.065	37.4007	9.350163	18.136	57.995	4.944	127.83
				Total	21	32.756	33.9317	7.404499	17.310	48.201	4.944	127.83
	Medium	**Undiversified**	**Sales growth**	Low	98	.17136	.272164	.027493	.11680	.22593	-.318	1.674
				High	95	.21832	.437304	.044866	.12924	.30740	-.204	3.128
				Total	193	.19448	.362770	.026113	.14297	.24598	-.318	3.128
			ROA	Low	99	.01695	.158984	.015978	-.01476	.04866	-.540	.525
				High	98	-.04664	.274830	.027762	-.10174	.00846	-1.539	.212
				Total	197	-.01469	.225908	.016095	-.04643	.01706	-1.539	.525
			EBIT margin	Low	99	.06191	.071853	.007221	.04758	.07624	-.137	.233
				High	98	.00835	.102997	.010404	-.01230	.02899	-.605	.127
				Total	197	.03526	.092479	.006589	.02227	.04826	-.605	.233
			P/E ratio	Low	50	19.617	12.1611	1.719839	16.161	23.073	7.204	60.73
				High	53	47.739	63.8589	8.771688	30.137	65.340	4.876	300.60
				Total	103	34.087	48.4714	4.776034	24.614	43.561	4.876	300.60
		Diversified	**Sales growth**	Low	38	.15346	.141045	.022881	.10710	.19982	-.103	.542
				High	30	.17295	.303189	.055355	.05974	.28617	-.223	1.415
				Total	68	.16206	.225542	.027351	.10747	.21665	-.223	1.415
			ROA	Low	39	.00670	.173961	.027856	-.04969	.06309	-.600	.170
				High	31	-.01160	.133376	.023955	-.06052	.03732	-.398	.136
				Total	70	-.00140	.156475	.018702	-.03871	.03591	-.600	.170

		EBIT margin	Low	39	.08726	.049802	.007975	.07111	.10340	-.003	.196
			High	31	.04818	.070613	.012682	.02228	.07408	-.082	.210
			Total	70	.06995	.062578	.007480	.05503	.08487	-.082	.210
		P/E ratio	Low	25	17.767	7.2000	1.439997	14.795	20.739	7.000	37.87
			High	18	26.820	35.0430	8.259709	9.394	44.247	6.169	159.40
			Total	43	21.556	23.3901	3.566963	14.358	28.755	6.169	159.40
Large	**Undiversified**	**Sales growth**	Low	90	.12132	.240730	.025375	.07090	.17174	-.295	1.846
			High	59	.17830	.292944	.038138	.10196	.25464	-.185	1.688
			Total	149	.14388	.263176	.021560	.10128	.18649	-.295	1.846
		ROA	Low	91	.04031	.064487	.006760	.02688	.05374	-.350	.197
			High	59	.00479	.289756	.037723	-.07073	.08030	-2.075	.217
			Total	150	.02634	.188406	.015383	-.00406	.05673	-2.075	.217
		EBIT margin	Low	91	.05981	.030497	.003197	.05346	.06616	-.028	.119
			High	59	.04552	.063798	.008306	.02889	.06215	-.364	.173
			Total	150	.05419	.046853	.003826	.04663	.06175	-.364	.173
		P/E ratio	Low	75	28.411	60.8769	7.029454	14.404	42.417	2.081	530.13
			High	46	25.901	15.4080	2.271785	21.325	30.476	5.296	80.96
			Total	121	27.457	48.7431	4.431188	18.683	36.230	2.081	530.13
	Diversified	**Sales growth**	Low	33	.17136	.307230	.053482	.06242	.28030	-.164	1.176
			High	46	.17066	.450903	.066482	.03676	.30456	-.066	3.040
			Total	79	.17095	.394994	.044440	.08248	.25943	-.164	3.040
		ROA	Low	33	.03080	.043266	.007532	.01546	.04614	-.081	.119
			High	46	.03035	.063199	.009318	.01158	.04912	-.281	.115
			Total	79	.03054	.055428	.006236	.01812	.04295	-.281	.119
		EBIT margin	Low	33	.07683	.046228	.008047	.06044	.09322	-.065	.157
			High	46	.05226	.039875	.005879	.04042	.06410	-.040	.176
			Total	79	.06252	.044077	.004959	.05265	.07240	-.065	.176
		P/E ratio	Low	25	29.700	26.8415	5.368296	18.620	40.780	9.302	124.96
			High	42	29.063	22.6264	3.491330	22.013	36.114	4.707	122.53
			Total	67	29.301	24.0856	2.942523	23.426	35.176	4.707	124.96

Table 45 Descriptives of the samples for SIC codes 20, 58 and 59, with different groups for the combinations of size and diversification segments.

ANOVA

SIC	Size (sales)	Diversification			Sum of Squares	df	Mean Square	F	Sig.
20	**Small**	**Undiversified**	**Sales growth**	Between Groups	.105	1	.105	1.387	.240
				Within Groups	14.988	197	.076		
				Total	15.093	198			
			ROA	Between Groups	.024	1	.024	.336	.563
				Within Groups	14.512	205	.071		
				Total	14.536	206			
			EBIT margin	Between Groups	.000	1	.000	.002	.963
				Within Groups	4.367	201	.022		
				Total	4.367	202			
			P/E ratio	Between Groups	496.559	1	496.559	.406	.525
				Within Groups	119817.721	98	1222.630		
				Total	120314.280	99			
		Diversified	**Sales growth**	Between Groups	.232	1	.232	1.568	.215
				Within Groups	9.037	61	.148		
				Total	9.269	62			
			ROA	Between Groups	.006	1	.006	.983	.325
				Within Groups	.393	63	.006		
				Total	.399	64			
			EBIT margin	Between Groups	.004	1	.004	.670	.416
				Within Groups	.403	61	.007		
				Total	.408	62			

			P/E ratio	Between Groups	828.680	1	828.680	1.041	.315
				Within Groups	27867.874	35	796.225		
				Total	28696.554	36			
	Medium	**Undiversified**	**Sales growth**	Between Groups	.094	1	.094	1.926	.168
				Within Groups	5.651	116	.049		
				Total	5.744	117			
			ROA	Between Groups	.024	1	.024	2.900	.091
				Within Groups	.997	119	.008		
				Total	1.021	120			
			EBIT margin	Between Groups	.023	1	.023	6.622	.011
				Within Groups	.420	119	.004		
				Total	.443	120			
			P/E ratio	Between Groups	584.283	1	584.283	.583	.447
				Within Groups	82227.806	82	1002.778		
				Total	82812.089	83			
		Diversified	**Sales growth**	Between Groups	.000	1	.000	.009	.927
				Within Groups	.870	22	.040		
				Total	.871	23			
			ROA	Between Groups	.000	1	.000	.003	.959
				Within Groups	.141	22	.006		
				Total	.141	23			
			EBIT margin	Between Groups	.000	1	.000	.110	.744
				Within Groups	.087	22	.004		
				Total	.087	23			
			P/E ratio	Between Groups	84.687	1	84.687	.620	.448
				Within Groups	1503.571	11	136.688		
				Total	1588.258	12			
	Large	**Undiversified**	**Sales growth**	Between Groups	.059	1	.059	1.803	.182
				Within Groups	3.129	96	.033		
				Total	3.188	97			
			ROA	Between Groups	.039	1	.039	8.282	.005
				Within Groups	.462	98	.005		
				Total	.501	99			
			EBIT margin	Between Groups	.138	1	.138	31.433	.000
				Within Groups	.430	98	.004		
				Total	.569	99			
			P/E ratio	Between Groups	311.074	1	311.074	.790	.377
				Within Groups	30704.554	78	393.648		
				Total	31015.628	79			
		Diversified	**Sales growth**	Between Groups	.245	1	.245	2.268	.136
				Within Groups	9.307	86	.108		
				Total	9.552	87			
			ROA	Between Groups	.047	1	.047	19.815	.000
				Within Groups	.207	87	.002		
				Total	.255	88			
			EBIT margin	Between Groups	.067	1	.067	38.726	.000
				Within Groups	.150	87	.002		
				Total	.217	88			
			P/E ratio	Between Groups	2903.534	1	2903.534	1.812	.184
				Within Groups	92960.536	58	1602.768		
				Total	95864.070	59			
58	**Small**	**Undiversified**	**Sales growth**	Between Groups	.461	1	.461	4.575	.033
				Within Groups	23.988	238	.101		
				Total	24.449	239			
			ROA	Between Groups	.073	1	.073	2.247	.135
				Within Groups	7.762	239	.032		
				Total	7.835	240			
			EBIT margin	Between Groups	.004	1	.004	.552	.458

			Within Groups	1.555	243	.006		
			Total	1.558	244			
		P/E ratio	Between Groups	5199.635	1	5199.635	8.799	.004
			Within Groups	69141.959	117	590.957		
			Total	74341.594	118			
	Diversified	Sales growth	Between Groups	.000	1	.000	.001	.971
			Within Groups	2.986	48	.062		
			Total	2.986	49			
		ROA	Between Groups	.227	1	.227	11.119	.002
			Within Groups	1.081	53	.020		
			Total	1.308	54			
		EBIT margin	Between Groups	.057	1	.057	10.184	.002
			Within Groups	.281	50	.006		
			Total	.339	51			
		P/E ratio	Between Groups	2662.133	1	2662.133	12.555	.001
			Within Groups	5725.059	27	212.039		
			Total	8387.192	28			
Medium	Undiversified	Sales growth	Between Groups	.073	1	.073	1.481	.225
			Within Groups	10.016	204	.049		
			Total	10.089	205			
		ROA	Between Groups	.003	1	.003	.070	.791
			Within Groups	8.012	206	.039		
			Total	8.014	207			
		EBIT margin	Between Groups	.007	1	.007	1.465	.227
			Within Groups	.958	206	.005		
			Total	.965	207			
		P/E ratio	Between Groups	877.665	1	877.665	.449	.504
			Within Groups	242286.764	124	1953.926		
			Total	243164.429	125			
	Diversified	Sales growth	Between Groups	.001	1	.001	.038	.846
			Within Groups	.594	26	.023		
			Total	.595	27			
		ROA	Between Groups	.005	1	.005	.570	.457
			Within Groups	.253	27	.009		
			Total	.259	28			
		EBIT margin	Between Groups	.004	1	.004	2.137	.155
			Within Groups	.054	27	.002		
			Total	.058	28			
		P/E ratio	Between Groups	10.420	1	10.420	.049	.829
			Within Groups	2120.303	10	212.030		
			Total	2130.723	11			
Large	Undiversified	Sales growth	Between Groups	.025	1	.025	1.488	.226
			Within Groups	1.295	76	.017		
			Total	1.320	77			
		ROA	Between Groups	.061	1	.061	3.428	.068
			Within Groups	1.359	76	.018		
			Total	1.420	77			
		EBIT margin	Between Groups	.117	1	.117	31.474	.000
			Within Groups	.283	76	.004		
			Total	.400	77			
		P/E ratio	Between Groups	226.402	1	226.402	2.294	.135
			Within Groups	6119.269	62	98.698		
			Total	6345.670	63			
	Diversified	Sales growth	Between Groups	.006	1	.006	.227	.638
			Within Groups	.581	24	.024		
			Total	.587	25			
		ROA	Between Groups	.000	1	.000	.036	.850
			Within Groups	.051	25	.002		

				Total	.051	26			
			EBIT margin	Between Groups	.000	1	.000	.047	.829
				Within Groups	.013	25	.001		
				Total	.013	26			
			P/E ratio	Between Groups	121.046	1	121.046	1.908	.186
				Within Groups	1015.296	16	63.456		
				Total	1136.342	17			
59	**Small**	**Undiversified**	**Sales growth**	Between Groups	1.498	1	1.498	4.692	.031
				Within Groups	73.431	230	.319		
				Total	74.929	231			
			ROA	Between Groups	1.818	1	1.818	10.989	.001
				Within Groups	42.015	254	.165		
				Total	43.833	255			
			EBIT margin	Between Groups	.830	1	.830	15.986	.000
				Within Groups	11.998	231	.052		
				Total	12.828	232			
			P/E ratio	Between Groups	2738.482	1	2738.482	1.074	.302
				Within Groups	270218.667	106	2549.233		
				Total	272957.149	107			
		Diversified	**Sales growth**	Between Groups	.627	1	.627	3.914	.053
				Within Groups	8.167	51	.160		
				Total	8.793	52			
			ROA	Between Groups	1.147	1	1.147	4.743	.033
				Within Groups	14.024	58	.242		
				Total	15.171	59			
			EBIT margin	Between Groups	.004	1	.004	.229	.634
				Within Groups	.903	50	.018		
				Total	.907	51			
			P/E ratio	Between Groups	1894.297	1	1894.297	1.703	.207
				Within Groups	21132.875	19	1112.257		
				Total	23027.172	20			
	Medium	**Undiversified**	**Sales growth**	Between Groups	.106	1	.106	.807	.370
				Within Groups	25.161	191	.132		
				Total	25.268	192			
			ROA	Between Groups	.199	1	.199	3.961	.048
				Within Groups	9.804	195	.050		
				Total	10.003	196			
			EBIT margin	Between Groups	.141	1	.141	17.950	.000
				Within Groups	1.535	195	.008		
				Total	1.676	196			
			P/E ratio	Between Groups	20346.788	1	20346.788	9.371	.003
				Within Groups	219300.303	101	2171.290		
				Total	239647.091	102			
		Diversified	**Sales growth**	Between Groups	.006	1	.006	.124	.726
				Within Groups	3.402	66	.052		
				Total	3.408	67			
			ROA	Between Groups	.006	1	.006	.234	.630
				Within Groups	1.684	68	.025		
				Total	1.689	69			
			EBIT margin	Between Groups	.026	1	.026	7.354	.008
				Within Groups	.244	68	.004		
				Total	.270	69			
			P/E ratio	Between Groups	857.808	1	857.808	1.590	.214
				Within Groups	22120.331	41	539.520		
				Total	22978.139	42			
	Large	**Undiversified**	**Sales growth**	Between Groups	.116	1	.116	1.678	.197
				Within Groups	10.135	147	.069		
				Total	10.251	148			

		ROA	Between Groups	.045	1	.045	1.275	.261
			Within Groups	5.244	148	.035		
			Total	5.289	149			
		EBIT margin	Between Groups	.007	1	.007	3.381	.068
			Within Groups	.320	148	.002		
			Total	.327	149			
		P/E ratio	Between Groups	179.675	1	179.675	.075	.785
			Within Groups	284926.664	119	2394.342		
			Total	285106.339	120			
	Diversified	**Sales growth**	Between Groups	.000	1	.000	.000	.994
			Within Groups	12.170	77	.158		
			Total	12.170	78			
		ROA	Between Groups	.000	1	.000	.001	.972
			Within Groups	.240	77	.003		
			Total	.240	78			
		EBIT margin	Between Groups	.012	1	.012	6.385	.014
			Within Groups	.140	77	.002		
			Total	.152	78			
		P/E ratio	Between Groups	6.350	1	6.350	.011	.918
			Within Groups	38281.279	65	588.943		
			Total	38287.629	66			

Table 46 ANOVA details for the samples for SIC codes 20, 58 and 59, with different groups for the combinations of size and diversification segments. The suffixes *, ** and *** indicate that differences between the high and low coherence groups are significant at the 0.10, .05 and .01 significance level respectively.

Test of Homogeneity of Variances							
SIC	**Size (sales)**	**Diversification**		Levene Statistic	df1	df2	Sig.
20	**Small**	**Undiversified**	**Sales growth**	.730	1	197	.394
			ROA	1.024	1	205	.313
			EBIT margin	3.835	1	201	.052
			P/E ratio	.682	1	98	.411
		Diversified	**Sales growth**	1.931	1	61	.170
			ROA	.869	1	63	.355
			EBIT margin	.238	1	61	.628
			P/E ratio	4.347	1	35	.044
	Medium	**Undiversified**	**Sales growth**	.100	1	116	.752
			ROA	3.713	1	119	.056
			EBIT margin	28.611	1	119	.000*
			P/E ratio	.555	1	82	.458
		Diversified	**Sales growth**	1.112	1	22	.303
			ROA	.836	1	22	.371
			EBIT margin	2.271	1	22	.146
			P/E ratio	.164	1	11	.694
	Large	**Undiversified**	**Sales growth**	1.129	1	96	.291
			ROA	.001	1	98	.982
			EBIT margin	6.934	1	98	.010*
			P/E ratio	.407	1	78	.526
		Diversified	**Sales growth**	4.070	1	86	.047*
			ROA	16.192	1	87	.000*
			EBIT margin	1.695	1	87	.196
			P/E ratio	7.943	1	58	.007*
58	**Small**	**Undiversified**	**Sales growth**	4.901	1	238	.028
			ROA	1.628	1	239	.203
			EBIT margin	2.398	1	243	.123
			P/E ratio	7.727	1	117	.006*
		Diversified	**Sales growth**	.198	1	48	.658
			ROA	12.166	1	53	.001*
			EBIT margin	67.942	1	50	.000*

			P/E ratio	21.377	1	27	.000*
	Medium	**Undiversified**	**Sales growth**	.254	1	204	.615
			ROA	.184	1	206	.668
			EBIT margin	5.233	1	206	.023
			P/E ratio	.132	1	124	.717
		Diversified	**Sales growth**	1.534	1	26	.227
			ROA	2.552	1	27	.122
			EBIT margin	1.408	1	27	.246
			P/E ratio	1.821	1	10	.207
	Large	**Undiversified**	**Sales growth**	2.443	1	76	.122
			ROA	8.430	1	76	.005*
			EBIT margin	8.864	1	76	.004*
			P/E ratio	.027	1	62	.871
		Diversified	**Sales growth**	5.013	1	24	.035
			ROA	17.931	1	25	.000*
			EBIT margin	3.616	1	25	.069
			P/E ratio	4.399	1	16	.052
59	**Small**	**Undiversified**	**Sales growth**	13.026	1	230	.000*
			ROA	9.304	1	254	.003*
			EBIT margin	19.862	1	231	.000*
			P/E ratio	3.688	1	106	.058
		Diversified	**Sales growth**	.048	1	51	.828
			ROA	12.773	1	58	.001*
			EBIT margin	.550	1	50	.462
			P/E ratio	7.777	1	19	.012*
	Medium	**Undiversified**	**Sales growth**	4.298	1	191	.040
			ROA	2.794	1	195	.096
			EBIT margin	1.060	1	195	.305
			P/E ratio	20.568	1	101	.000*
		Diversified	**Sales growth**	2.896	1	66	.094
			ROA	.232	1	68	.632
			EBIT margin	3.134	1	68	.081
			P/E ratio	5.791	1	41	.021*
	Large	**Undiversified**	**Sales growth**	2.714	1	147	.102
			ROA	5.663	1	148	.019
			EBIT margin	.975	1	148	.325
			P/E ratio	.986	1	119	.323
		Diversified	**Sales growth**	.079	1	77	.779
			ROA	.050	1	77	.823
			EBIT margin	1.543	1	77	.218
			P/E ratio	.099	1	65	.754

Table 47 Levene's test of homogeneity of variances for the samples for SIC codes 20, 58 and 59, with different groups for the combinations of size and diversification segments. Suffix * indicates that the variances between the two coherence groups are unequal at the 0.05 significance level. This means that the ANOVA assumption of equal variances is not met and this can influence the ANOVA results.

Independent Samples Kruskal-Wallis Test (Non-parametric One-way ANOVA)				
SIC	**Size (sales)**	**Diversification**	**Premium**	**Sig.**
20	**Small**	**Undiversified**	**Sales growth**	.006
			ROA	.058
			EBIT margin	.207
			P/E ratio	.759
		Diversified	**Sales growth**	.092
			ROA	.082
			EBIT margin	.749
			P/E ratio	.647

	Medium	Undiversified	**Sales growth**	.036
			ROA	.003
			EBIT margin	.089
			P/E ratio	.312
		Diversified	**Sales growth**	.739
			ROA	.594
			EBIT margin	.594
			P/E ratio	.310
	Large	Undiversified	**Sales growth**	.659
			ROA	.001
			EBIT margin	.000
			P/E ratio	.272
		Diversified	**Sales growth**	.119
			ROA	.000
			EBIT margin	.000
			P/E ratio	.426
58	Small	Undiversified	**Sales growth**	.016
			ROA	.209
			EBIT margin	.892
			P/E ratio	.003
		Diversified	**Sales growth**	.683
			ROA	.014
			EBIT margin	.578
			P/E ratio	.036
	Medium	Undiversified	**Sales growth**	.005
			ROA	.005
			EBIT margin	.196
			P/E ratio	.145
		Diversified	**Sales growth**	.476
			ROA	.661
			EBIT margin	.495
			P/E ratio	.926
	Large	Undiversified	**Sales growth**	.041
			ROA	.018
			EBIT margin	.000
			P/E ratio	.096
		Diversified	**Sales growth**	.506
			ROA	.699
			EBIT margin	.658
			P/E ratio	.173
59	Small	Undiversified	**Sales growth**	.521
			ROA	.000
			EBIT margin	.000
			P/E ratio	.900
		Diversified	**Sales growth**	.024
			ROA	.084
			EBIT margin	.631
			P/E ratio	.509
	Medium	Undiversified	**Sales growth**	.908
			ROA	.158
			EBIT margin	.000
			P/E ratio	.005
		Diversified	**Sales growth**	.639
			ROA	.168
			EBIT margin	.004
			P/E ratio	1.000
	Large	Undiversified	**Sales growth**	.223
			ROA	.855

Diversified	**EBIT margin**	.016
	P/E ratio	.063
	Sales growth	.921
	ROA	.409
	EBIT margin	.003
	P/E ratio	.826

Table 48 Kruskal-Wallis Test (Non-parametric One-way ANOVA) for the samples for SIC codes 20, 58 and 59, with different groups for the combinations of size and diversification segments. The suffixes *, ** and *** indicate that differences between the high and low coherence groups are significant at the 0.10, .05 and .01 significance level respectively.

Glossary

Abbreviation	Explanation
10-K	Yearly SEC filing, obligatory for publicly traded companies active in the USA.
ANOVA	ANalysis Of Variance, determines differences between the means of two or more samples
APQC	American Productivity and Quality Center
CEO	Chief Executive Officer
DC	Dynamic Capabilities
EBIT	Earnings Before Interest and Taxes
EBITDA	Earnings Before Interest and Taxes Depreciation and Amortization
EDGAR	Electronic Data-Gathering, Analysis and Retrieval system
EE	Evolutionary Economics
FamaFrench-12	A classification into twelve industries, invented by Eugene Fama and Kenneth French
FamaFrench-48	A classification into 48 industries, invented by Eugene Fama and Kenneth French
IO	Industrial Organisation
MD&A	Management Discussion and Analysis, Item 7 of the SEC 10-K filing
NI	Net Income
OLS	Ordinary Least Squares
P/E	Price/Earnings = Price/Net Income
RBV	Resource-Based View
ROA	Return on assets
ROI	Return on investment
SEC	US Securities and Exchange Commission
SIC	Standard Industry Classification