QUALITY ASSURANCE IN CORPORATE FINANCIAL PLANNING

– A PROCESS- AND DATA-DRIVEN PERSPECTIVE –

Zur Erlangung des akademischen Grades eines
Doktors der Wirtschaftswissenschaften

(Dr. rer. pol.)

von der Fakultät für
Wirtschaftswissenschaften
am Karlsruher Institut für Technologie (KIT)

genehmigte

DISSERTATION

von

Dipl.-Math. oec. Jochen Martin

Tag der mündlichen Prüfung: 15.11.2012
Referent: Prof. Dr. Thomas Setzer
Korreferent: Prof. Dr. Hansjörg Fromm

Karlsruhe, 2012

Summary

The financial crisis has kept the world busy from 2007 to the present. The resulting decrease of confidence in the creditworthiness of most enterprises, banks, and industrial corporations, has made it painfully obvious that financial planning requires particular attention. Of course, it has been recognized long before the crisis that a precise forecast of business figures like sales, production, and investments is essential. Faulty hedge-positions resulting from inefficient exposure planning can increase hedging costs. Similar costly effects can result from imprecise liquidity planning, even if insolvency is avoided: high carry costs for liquidity are caused by high risk premium along with low interest on deposits, nearly independent of a company's structure and size.

Challenges for optimizing corporate financial planning data integration preceding risk management are even greater in global companies due to distributed and heterogeneous data generation processes. Research and contribution in this thesis address process and data related challenges and comprise of two parts: the first *process-driven* part evaluates the effect of corporate financial planning redesign based on an appropriate business process redesign model for multinational enterprises. Thereby, the focus is on a flexible execution structure and the three process related quality dimensions *timeliness*, *completeness*, and *consistency*. The second *data-driven* research part comes up with new quality metrics for financial planning data and their benchmarking against the forth quality dimension *accuracy*. Besides a detailed documentation of the status quo, the informational content of accuracy is queried and the idea is introduced to combine this metric with additional statistical measures and expert knowledge to support financial controllers. Both research parts are investigated through evaluation studies based on empirical data. The implementation of the first redesign step and a documentation of process runtime before and after redesign show a significant increase in *timeliness*, *completeness*, and *consistency*. Moreover, the conformity with Benford's Law is introduced as new quality metric based on detailed benchmarking analyses against the well-accepted quality metric *accuracy*. One major advantage of this newly introduced metric is the ability to indicate planning data quality at the moment of data generation, since the corresponding actual data is not required. Finally, previously unknown business insights are derived through the combination of new and old quality metrics and a stepwise in-depth investigation of planned and actual values. With the evaluations conducted throughout this thesis the overall goal of research with practical impact is achieved and promising fields of future work are identified.

Acknowledgements

I am indebted to many people for their support and collaboration in the course of my Ph.D. research. Foremost, I thank my supervisor, Prof. Dr. Thomas Setzer for his strong support and valuable discussion. For pushing my research into the right direction and providing important contributions through the last year of this thesis. Prof. Dr. Hansjörg Fromm for pleasant discussions and for co-advising this thesis. I would like to thank Prof. Dr. Christof Weinhardt for his continuous support, for the opportunity to start this thesis, and for serving the board of examiners. Prof. Dr. Martin Ruckes for completing the board of examiners. I specially thank Dr. Tobias Conte for his guidance and the valuable contributions he gave to this thesis.

I thank Bayer AG, Corporate Finance for the funding of this work and Alexander Burck for supporting all empirical studies as well as their publication. Additional thanks go to Prof. Dr. Martin Glaum, Kati Schuerer and Thomas Strehlow for fruitful discussion and cooperation.

I would like to express my gratitude to the team of the research group on Information and Market Engineering at the Institute of Information Systems and Marketing (IISM) as well as to the team of the research division on Information Process Engineering (IPE) at the FZI Research Center for Information Technology.

Special thanks go to:
Dr. Mercè Müller-Gorchs, Dr. Benjamin Blau, Prof. Dr. Stefan Seifert, and Stefanie Betz for their support in the early stage of this research and their seminal ideas.
Athanasios Mazarakis and Dr. Florian Teschner for the challenging discussion on statistics and their contribution to this work.
Nico Roedder and Dr. Florian Teschner for proofreading and constructive criticism.
My loyal companions Dr. Tobias Conte, Steffen Haak, Rico Knapper, and Athanasios Mazarakis for contributing valuable input to my research, and most important, for providing a relaxed and fun environment that eased working a lot.

Above all, I am indebted to my parents, who supported and encouraged me throughout the eventful time of my studies. My sister Judith who also had the questionable fortune of proofreading this work. Finally, I am particularly grateful to my fiancée Bianca for her unreserved love, patience and cheering me up when needed.

Jochen Martin

Contents

List of Figures

List of Tables

List of Abbreviations

Part I.

Foundations

Chapter 1.

Introduction

The financial crisis has kept the world busy from 2007 to the present day. The resulting decrease of confidence in the creditworthiness of most enterprises, banks, as well as industrial corporations, has made it painfully obvious that financial planning requires particular attention. Indeed, the effort conducted to financial planning processes in almost any kind of company raised, not least because it is a highly knowledge-intensive task. Of course, it has been recognized long before the crisis that a precise forecast of business figures like sales, production, and investments is essential (Kim et al., 1998; Graham and Harvey, 2001). Actually, 20% of enterprise insolvencies in Germany could have been been avoided through solid liquidity forecast and management (Schneider-Frisse, 2009). Moreover, the avoidance of faulty hedge-positions resulting from inefficient exposure planning can significantly decrease hedging costs.

Similar costly effects can result from imprecise liquidity planning, even if insolvency is avoided: for instance, as one result of the suffering confidence in the creditworthiness of the participants in capital markets the risk premiums for liquidity procurement increased. The fact that even large enterprises were hit by the crisis resulted in an increased awareness for the striking importance of a high credit rating to be able to access the capital market at reasonable costs. In addition, the governmental rescue measures induced cheap liquidity which accounted for low interest incomes and, hence, increased the cost of carry for companies that produce a liquidity surplus. The high carry costs for liquidity resulting from risk premium along with low interest on deposits made improper liquidity planning very expensive, nearly independent of a company's structure and size.

1.1. Motivation

The omnipresent effects of the financial crises led to a significantly increased awareness for the importance of proper forecasting as it was documented, for instance, in the "Treasury & Risk – 2010 Strategic Treasury Survey"[1]: 50% of the asked senior financial executives rated liquidity management the most promising area to increase efficiency in the next years (Anonymous, 2010). One of the key tasks to increase efficiency besides the commonly known quality assurance measures is process integration. Especially in global enterprises, vast heterogeneity in the data generation raises serious risks during this integration (Innig, 2009). Furthermore, increased efficiency is of utmost importance, since rolling financial planning with refinements in intervals of a few months and numbers on a very detailed level is required in highly volatile and insecure times (Innig, 2009). To support this effort in risk management, IT providers developed and enhanced integrative software solutions for corporate planning over the last years (Niemann, 2009).

For global companies, it is even more challenging to compose a well-founded planning (Goedhart and Spronk, 1995): in the case of central currency-specific liquidity planning, decentralized planning processes have to be coordinated within local partitions and internal transactions between these partitions have to be monitored to ensure a proper and consistent overall financial planning. Yet, global enterprises are subject to continuous growth and change through M&A measures and spin-offs. The resulting organizational heterogeneity is intensified by a multitude of business cul-

[1] www.treasuryandrisk.com

tures in the respective countries. Moreover, differences in enterprise size and hierarchical position in the group as a whole lead to numerous applied software solutions, each of which may work on individual data formats. Consequently, standardized software solutions supporting crucial planning tasks are oftentimes not applicable. Nevertheless, better technology is required in cash flow forecasting, budgeting, and planning (Anonymous, 2010). Especially, the process of data transmission and validation accompanied by an intensive communication between local and global management is important to assure high data input quality for risk management. Moreover, regardless of all challenges arising within complex company structures, the requirements placed on sensitive processes like financial management continuously increase (Lu et al., 2008).

To cope with these challenges in global companies, corporate financial portals have turned out to be an efficient measure to enhance the process of centralized liquidity risk management (Vo and Elsner, 2007). Today, such Information Systems (IS) are included into service-oriented architectures. In this vein, IT-based business intelligence services can be offered to support planning activities, such as market-based prediction services, services to detect complex events as a sequence of defined activities, or decision support services. During the implementation of such an information system, major challenges arise from high heterogeneity of applications, business processes, and employees within multinational enterprises as described above. In corporate financial planning these historically grown structures result in a knowledge intensive data transmission and communication process. Although such processes can be structured on a highly aggregated level (cp. Figure 3.1), numerous individual tasks, for instance, data validation and communication, lead to strongly individualized process workflows. The resulting importance of human participants in processes again impedes their improvement due to semi-structured process parts and potentially limited willingness to change and to abandon known structures and workflows (Riege, 2005). To cope with these challenges and to assure process quality at large, an efficient and customer friendly organizational structure and data consistency are equally important (Schultz, 1992). Up to now, literature has brought about several approaches to redesign and improve complex processes, however, most articles include either strong process restrictions or complex redesign procedures which are hardly applicable in practice.

The above-mentioned decision support services, especially if large amounts of underlying data are supposed to be processed for the planning task, are usually based on data mining methods. Thereby, detecting patterns in huge data sets can be tedious, for instance, since it requires a variety of upstream and downstream efforts and may be different for each data set considered (Witten et al., 2005). Moreover, quality of financial planning is usually quantified by its outcome using accepted ex-post metrics such as planning accuracy or alternative derivatives of plan versus actual deviances – planning errors. However, cash flow positions are often planned months or even years before they actually take place, with frequent revisions of their monetary quantification based on additional knowledge and adapted expectations. This work considers the sequence of financial planning data predicting the same ultimate cash flow (item) at different points in time as a documentation of the financial planning process. The goal of such effort is, first, to identify erroneous or suspicious planning data and revisions that will likely result in huge planning errors. Second, the determination of flawed planning processes allows for more profound root cause analysis of poor planning accuracy. Unfortunately, controllers today have little guidance on how to assess planning processes' quality at process runtime. This is particularly true because of the complex data structure in financial planning processes accompanied by often unknown assumptions and dynamics.

Now, what if it was possible to find properties valid for a variety of data sets, independent of the respective industry, task, or company? Benford's Law, as shown by Benford (1938), provides highly interesting insights into the structure of empirical data: naively thought it seems obvious that the digits of numbers in the decimal system are equally distributed. Yet, Benford showed for several kinds of empirically gathered numbers that the leading digits as well as the digits in second position occur with distinct probabilities which clearly differ from an equal distribution. If the digit distribution in high-quality financial planning data in fact follows Benford's Law, certain rounding behaviours or the creation of duplicate numbers via copy-and-pasting should distort the digit distribution. However, conformity to Benford's Law can be seen as a verisimilar of reported financial planning data but does not exhibit temporal structures present in planning data (repetitive patterns in consecutive revisions and planning errors). In order to analyse and describe data dependencies over whole planning processes, this work proposes a second metric: weak-form planning efficiency. The term weak-form planning efficiency is an adoption of weak-form forecast efficiency as it has been introduced by Nordhaus (1987) and relates to the statistical assumption of independent planning errors and revisions in financial planning data.

Overall, the conducted effort in all fields aims at an increased quality of the data input for risk management. Batini et al. (2009) show in their work, based on an excellent overview of existing approaches, that assessing and improving data quality in complex processes like financial planning requires differentiated procedures covering a wide range of goals. Yet, three stages are essential for quality optimization approaches: *(i)* the documentation of the organizational status quo in **state reconstruction**, *(ii)* measuring the data quality in **assessment/measurement**, and (iii) the **improvement**. As easily can be seen in these three stages, the documentation of the status quo is a key challenge that has to be present in quality improvement approaches. The four core-dimensions of quality documentation are *completeness*, *consistency*, *timeliness*, and *accuracy* (Batini et al., 2009; Wand and Wang, 1996). All data quality dimensions and their measurement are presented and discussed in detail in Section 2.2.

In particular, the focus of this thesis is to optimize process structure and output data quality of the data validation and aggregation process preceding the liquidity and foreign exchange risk management in a multinational enterprise (Singhvi, 1972; Martin et al., 2012). Consequently, research and contribution address process and data related challenges and comprise of two parts in analogy to Batini et al. (2009): the first *process-driven* part evaluates the effect of corporate financial planning redesign based on an appropriate business process redesign model for multinational enterprises. Thereby, the focus is on a flexible execution structure and the three process related quality dimensions. Based on the optimized organizational structure, the second *data-driven* research part comes up with new quality metrics for financial planning data and their benchmarking against the forth quality dimension data *accuracy*. Besides a detailed documentation of the status quo as requested by Batini et al. (2009), the informational content of accuracy and its role as a single, dominant quality metric is queried. As an alternative, the idea is introduced to combine this metric with additional statistical measures and expert knowledge to support financial controllers. One major advantage of the newly introduced metric will be the ability to indicate planning data quality at the moment of data generation, since the corresponding actual data is not required. Finally, previously unknown business insights will be derived through the combination of new and old quality metrics and a stepwise in-depth investigation of planned and actual values. In doing so, it should be possible to provide valuable support for the decision, which data samples should be further investigated and what kind of delivered planning data is likely to result in huge final planning errors.

1.2. Research Questions

Quality assurance in financial planning is a multi-step procedure. As above-introduced, the first step in this procedure is the process optimization which is necessary to directly increase data quality and to create space for complex quality assurance measures through reduced execution time. The special challenges arising from the financial planning in multinational enterprises result in the necessity for an innovative business process redesign framework. Although numerous approaches to redesign and improve complex processes exist in literature, an approach applicable in practice without strong process restrictions is missing. This gap is addressed in this work by investigating the first general research question (RQ):

RQ 1 – INNOVATIVE BUSINESS PROCESS REDESIGN –
Can a redesign model be derived that increases data quality and similarly offers standardization and flexibility to assure practical relevance?

Since quality assurance measures are very time consuming, the focus of such a model should be put on reduced time effort, which is a quality dimension itself (cp. 2.2). Nevertheless, as will be shown in the evaluation, the redesign model also affects the further quality dimensions completeness and consistency in parallel.

The second step in the procedure of quality assurance in financial planning has a clear data perspective: the extraction of information from historical planning data about the individual planning behaviour of entities. The empirical data sample underlying the analyses contains planning data at our industrial partner over a multiple year time period enriched with the corresponding actual values per planned value. The goal of such procedures is to transform existing knowledge into decision support services. In this vein, the approaches presented in this work are two-fold: *(i)* introducing existing quality indicators from other fields to financial planning with the goal of an ex-ante quality metric (actual data is not required), and *(ii)* examining common patterns in planning and actual financial data. As introduced earlier, Benford's Law provides highly interesting insights into the structure of empirical data. Conformity to Benford's Law does not exhibit temporal structures present in planning data (repetitive patterns in consecutive revisions and planning errors). Consequently, the investigation of *(i)* also refers to weak planning efficiency derived from weak forecast efficiency as introduced by Nordhaus (1987) and aims at answering the second main research question:

RQ 2 – EX-ANTE QUALITY METRIC –
Can existing quality metrics be introduced to financial planning data to generate an ex-ante quality indication?

Furthermore, as above-mentioned, it is of interest to identify patterns beyond planning data in the relationship to the corresponding actual data. The findings provided by the evaluation of **RQ 2** enable us to further investigate the special structure of the basic data in depth. Along with extensive expert knowledge about company characteristics, it is possible to address the research question:

RQ 3 – BUSINESS INSIGHTS –
Can compositions of actual and planning data provide business insights?

To specify these three overall questions and to structure their examination, the following three Sections 1.2.1 to 1.2.3 present sub-questions for each of the above-mentioned high-level research questions.

1.2.1. RQ 1 – Innovative Business Process Redesign –

Design and evaluation of the framework requested in **RQ 1** raises numerous challenges. The first sub-research question below addresses the design itself. To evaluate the appropriateness of the resulting flexible objective-based process redesign model, we applied it in a renowned, globally operating large company acting in the chemical and pharmaceutical sector. As introduced earlier, we focus on three quality dimensions, whereby *timeliness* is addressed in **RQ 1.2**, and *completeness* along with *consistency* are combined in **RQ 1.3**.

RQ 1.1 – INNOVATIVE BUSINESS PROCESS REDESIGN –
How should a theoretically based process redesign model that combines standardization and flexibility be designed to assure practical relevance?

RQ 1.2 – INNOVATIVE BUSINESS PROCESS REDESIGN –
Does the objective-based process redesign increase the data quality dimension timeliness in practice?

RQ 1.3 – INNOVATIVE BUSINESS PROCESS REDESIGN –
Does the objective-based process redesign increase the data quality dimensions completeness and consistency in practice?

The empirical evidence regarding the sub-research questions **RQ 1.2** and **RQ 1.3** is provided by an implementation of a redesigned process: at our industry partner, the redesigned processes are a part of daily routines since June 2010. Driven by the fact that process redesign literature is mostly of theoretical nature, the real world evaluation of the redesign model is one of the core contributions of this work.

1.2.2. RQ 2 – Ex-Ante Quality Metric –

In order to answer **RQ 2**, we apply Benford's Law and weak-form planning efficiency to multi-year financial planning data from over hundred subsidiaries. If it is possible to show that the digit distribution in financial planning data does in fact follow Benford's Law, data characteristics can be studied based on the conformity with Benford's Law. The results can be combined with additional statistical evaluations and expert knowledge to eventually support financial planning managers in their decision which data samples to further investigate. For instance, certain rounding behaviours or the creation of duplicate numbers via copy-and-pasting should distort the digit distribution. In order to analyse and describe data dependencies over whole planning processes, the second metric weak planning efficiency is applied to the same data sample. The term weak planning efficiency relates to the statistical assumption of independent planning errors and revisions in corporate financial planning data. Consequently, the following three sub-research questions were investigated:

RQ 2.1 – EX-ANTE QUALITY METRIC –
Does financial planning data follow Benford's Law?

RQ 2.2 – EX-ANTE QUALITY METRIC –
Do company characteristics impact the conformity of financial planning data with Benford's Law?

RQ 2.3 – EX-ANTE QUALITY METRIC –
Do company characteristics impact the weak planning efficiency of financial planning data?

Finally, it is of particular importance for the validity of Benford's Law based decision support to evaluate the quality indication provided by a conformity with Benford's Law. Moreover, quality assessment beyond data accuracy, i.e. the difference between plan and actual values, is crucial to overcome the lack of missing actual data at the moment of planning data generation. To elaborate whether perceived data quality improvements are reflected in real data accuracy and to evaluate the validity of Benford's Law as a quality indicator the following two final sub-research questions to **RQ 2** are examined:

RQ 2.4 – EX-ANTE QUALITY METRIC –
Is data accuracy conform with perceived data quality?

RQ 2.5 – EX-ANTE QUALITY METRIC –
Does Benford's Law in financial planning data assess data quality?

Through this evaluation the validity of Benford's Law would be based on its ability to reflect data characteristics previously discovered based on data accuracy. With positive findings it would be possible to start developing a decision support service based on digital analyses. However, for the appropriate design of such services the extraction of additional entity specific knowledge is of significant importance.

1.2.3. RQ 3 – Business Insights –

The first step to gain such entity specific information are data classifications like, for instance, by business line (cp. **RQ 2.2/2.3**). Beyond a business line specific conformity with Benford's Law, the evaluation of further patterns in planning and actual data promises valuable insights. In that way, potential differences between Benford, weak planning efficiency, and accuracy indications can be the starting point for an in depth investigation and the foundation for managerial impact. The top-down character of our approach through the examination on group and entity level is expressed by the following three sub-research questions for **RQ 3**:

RQ 3.1 – BUSINESS INSIGHTS –
Can subgroup planning data accuracy be related to external factors depending on the respective line of business?

RQ 3.2 – BUSINESS INSIGHTS –
Can conformity with Benford's Law or weak planning efficiency provide planning quality indications beyond data accuracy?

RQ 3.3 – BUSINESS INSIGHTS –
Can entity specific recommendation be derived based on Benford's Law, weak planning efficiency, and data accuracy evaluation?

Summing up, the analyses performed in this work are a multi-step approach diving deep into the process and data structures of corporate financial planning to extract systematic planning behaviour. The answers to the three research questions **RQ 1** to **RQ 3** in the Chapters 3 to 6 provide theoretically based investigations with a focus on their practical impact and on concrete recommendations for data adaptations.

1.3. Structure and Research Development

The earlier part of this chapter already gave a broad motivation onto why an improvement in financial planning is necessary at all and how it can be achieved. Beyond that, Section 1.2 described the research question and their interdependencies. The structure of the research questions is completely in line with the structure of the entire thesis: starting with the foundations in *Part I.*, *Part II.* (**RQ 1**) and *Part III.* (**RQ 2** and **RQ 3**) present the detailed investigation results before *Part IV.* concludes with a summary of the contribution and the future research topics. A high-level illustration of this structure can be found in Figure 1.1.

The research that resulted in this thesis was part of a research cooperation with a globally acting large enterprise in the pharmaceutical and chemical industry. Our *industry partner* can be seen as an archetypical multinational enterprise with subsidiaries

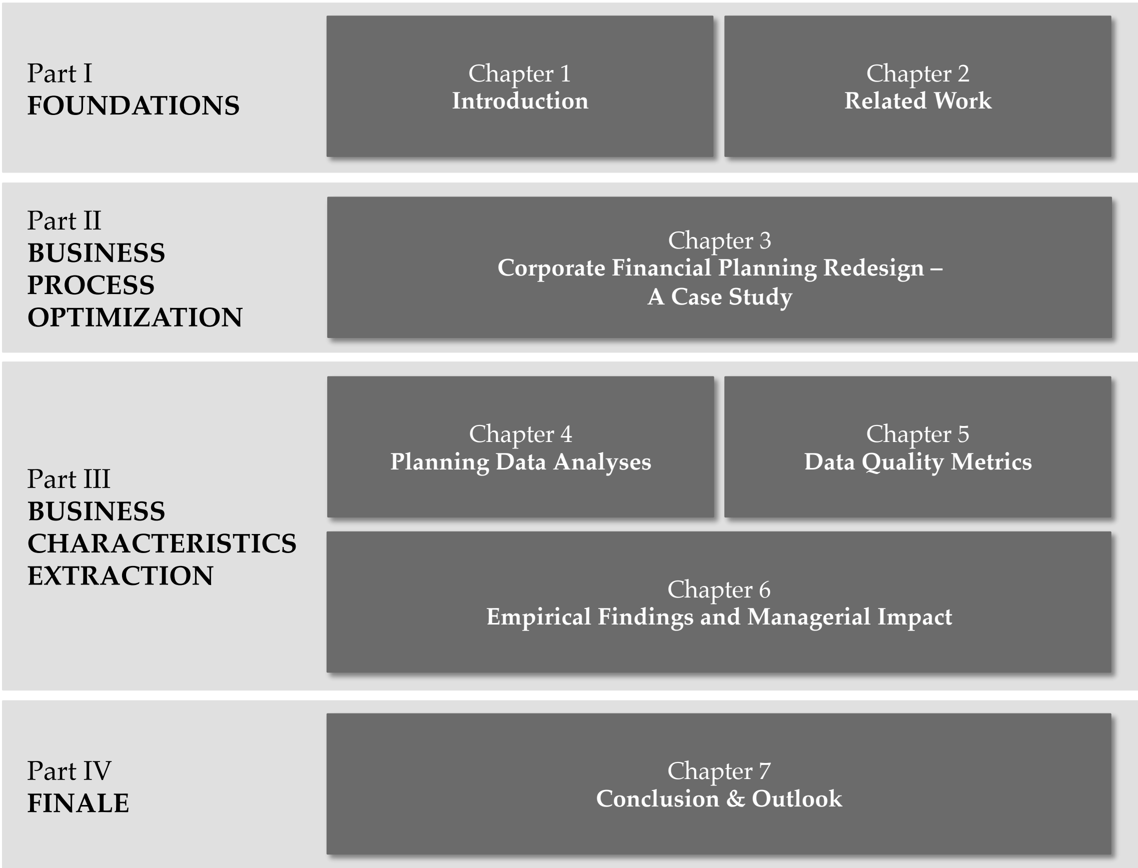

Figure 1.1.: Structure of this thesis.

spread all over the world. This implies a decentralised data generation. Moreover, continuous growth leads to a waste heterogeneity of employees and applied systems. Even more important, with planning numbers delivered, for instance, on the currency and partner level every three months, the organisational structure of the financial planning process exactly fulfil the previously mentioned detail and frequency requirements (Innig, 2009; Saroney III, 2005). Our experts within the enterprise are located in the corporate financial planning department within the holding and participate in the redesigned process as knowledge workers and responsible managers. The overall goal of this work was to perform profound research with practical impact making use of the evaluation opportunities offered within the cooperation. The conducted scientific effort lead to numerous publications in conference proceeding, all of which were reviewed and presented in advance.

The following gives a brief overview of these publications as a part of the respective chapters. Throughout this work, "we" refers to both; the readers of this work and me and my co-authors. I would like to particularly thank Thomas Setzer, Christof Weinhardt, Tobias Conte, Athanasios Mazarakis, Florian Teschner, Rico Knapper, Stefanie Betz, Simon Caton, and Chris Gerhardt. A first article describing the challenges in redesigning financial planning processes in multinational enterprises was discussed at the Group Decision and Negotiation 2010 Conference (Martin and Blau, 2010). Moreover, the research agenda for this thesis, including a short version of the model and a description of the planned data analyses, was presented and widely discussed during the Doctoral Consortium of the 10^{th} International Conference on Wirtschaftsinformatik (Martin, 2011).

In more detail, Chapter 2 provides the related literature in the field of business process redesign, quality metrics and data exploration in general. Based upon this foundation, Chapter 3 starts with a specification of the application domain. Thereafter, Section 3.2, describes our design science based objectives-based process redesign. This redesign model, including a formal representation of the requirements and a qualitative evaluation, was published in the proceedings of the 16^{th} Americas Conference on Information Systems (Martin et al., 2011). The core contribution in the field of business process redesign is the model evaluation through a real-world application in Sections 3.3 to 3.5, containing case study, current implementation status, and finally results. Thereby, the case study at our industry partner is documented in the proceedings of the 19^{th} European Conference on Information Systems (Martin et al. (2011) in Section 3.3). The service oriented implementation approach supporting a correct and efficient implementation of the newly developed process, was presented at the 8^{th} International Conference on Service Computing (Martin et al. (2011) in Section 3.4). Beyond this realization documentation of the redesign model we also successfully submitted the results of the implementation and quantitative evaluation to the 24^{th} International Conference on Advanced Information Systems Engineering (Martin and Conte (2012) in Section 3.5).

In parallel, we started implementing the first data analyses based on Benford's Law that are introduced in *Part III.*. The first results indicating the conformity of financial planning data with Benford's Law were published in the proceedings of the 20^{th} European Conference on Information Systems (Martin et al. (2012) in Section 4.2). Encouraged by the positive results and feedback, we performed in depth analyses regarding the digit distribution and weak planning efficiency. A comparison of both metrics was presented and discussed at the SRII Global Conference 2012 (Martin et al. (2012) in Section 4.3). In Chapter 5 similarities between perceived data quality and data accuracy are evaluated. Yet, one of the most important results are achieved in Section 5.2, where we establish the conformity with Benford's Law as data quality indicator. Based upon these findings, Chapter 6 presents astonishing insights into the data structure gained through a combination of the introduced quality metrics. Until now, these interesting results are unpublished, but a journal publication is in progress and we are currently working on the incorporation of the findings in a decision support service. Chapter 7 concludes with a contribution summary and the limitations of the performed research, pointing to future research topics.

According to the description above, the publications form the backbone of this thesis. Consequently, parts of them are included literally in the respective Chapters 3 and 4. Finally, some literally parts of the papers also found their way into introduction, methodology parts of the respective chapters, and outlook.

Chapter 2.

Related Work

The following section will lay the foundation regarding the theoretical background for this work. The introduction already gave a brief overview of the current effort for improvement undertaken in the domain of financial planning. Section 2.1 provides a broad overview of business process management literature necessary for the development of the redesign framework described in the following chapter. The quality measures presented in Section 2.2 allow us to assess the success of first the process redesign and second the data exploration analyses, for which the foundation is presented in Section 2.3.

2.1. Business Process Redesign

For the decision whether to talk about *reengineering* or *redesign* it is necessary to think about the scope of the restructuring effort. The scope of this work is to change the structure of one single process and therefore we use the term redesign according to Mansar and Reijers (2007), although there exists no explicit definition of the terms reengineering and redesign in literature (O'Neill and Sohal, 1999). Yet, reengineering is often associated with more drastic change programs (Mansar and Reijers, 2007). In general reengineering assumes a much broader scope than the specific focus of process redesign. Process redesign concentrates on the process *itself* in terms of its interdependent tasks and resources, while reengineering refers to all aspects of restructuring organization's processes, e.g. from change management to project management issues. Moreover, our business process definition follows Oberweis (1996), who describes a business process as "a set of manual, semi-automated or automated activities that are performed according to certain rules to achieve a particular business goal". With this strong emphasis on workflows within an enterprise rather than on products, this classification can be rated as workflow oriented process definition and is appropriate for enterprise systems application in this work (Davenport, 1993).

In the following section we characterize existing literature with respect to the required degree of standardization and point out the necessity for an alternative model. Thereby, the business process redesign model presented in Chapter 3 is designed to offer optimal support in semi-structured data delivery and interaction processes (for detailed motivating example refer to Section 3.1.1), however, it can be applied to semi-structured process in general as they are defined in Section 2.1.1. Of course, a redesign model able to cope with less structured processes can be applied to completely structured processes, too. Yet, as can be seen in the detailed overview of existing process redesign literature (Section 2.1.2), appropriate models already exist for this kind of processes.

2.1.1. Semi-Structured Processes

Generally, business processes can be distinguished by their level of structure. In this vein, Deiters (2000) distinguishes between structured, semi-structured, and completely unstructured business processes as follows: structured processes are applied in standardized scenarios and, therein, the sequence of tasks and business rules is predetermined and prescribed. In semi-structured processes, some tasks are not ordered at all and some of the rules may be modified or added later "on the fly". Hence, only parts of the sequence of tasks and the business rules are structured. Finally, unstructured processes do not have any repeatable patterns at all, are executed spontaneously, and are difficult to automate.

Löffeler et al. (1998) presented detailed evidence for the flexibility arising from an additional process type and its necessity. Their characterization is based on a metric consisting of the variables information base, co-operating partners, and solution path, previously identified by Picot and Reichwald (1984). To ease the identification of a process' structuredness, they introduce a classification framework. Although, such detailed definitions of semi-structured processes exist, they are not frequently considered in business redesign theory. This is particularly surprising, since semi-structured processes are omni-present in practice.

The data delivery and interaction process described in Section 3.1.1 follows a structured sequence of tasks on a high level (cp. Figure 3.1), however, the execution order of numerous tasks strongly depends on the individual preferences of the knowledge worker and may be modified on runtime. According to Deiters (2000), we consequently talk of semi-structured processes.

2.1.2. Business Process Redesign

To provide further insights into the characteristics of processes, lots of work has been performed on the analyses of processes' structuredness. Pentland (2003a) introduces sequential routines in business processes as a metric for the degree of standardization. Thereby, the identification of routine patterns allows for the definition of a lexicon per process and, hence, the comparison of different process parts. Furthermore, changes in the sequential execution of identified process patterns reflect either development or variety and, thus, a lower level of structuredness in the process (Pentland, 2003a). In an following benchmarking study, Pentland (2003b) shows that the variety of sequences rather expresses procedural knowledge and consequently derives a more generalized view than explicit task declarations in previously known measures like task variety. Rosenkranz et al. (2009) try to derive a unique pattern base to compare different workflows and to detect joint aspects as a foundation for a process standardization approach. Yet, their experiences in multiple case studies reveal complex challenges in redesigning business processes.

An additional shortcoming of recent literature beside the complexity of the introduced metrics for measurement and treatment of variety, is the focuses on either completely structured or unstructured processes. van der Aalst (2000), for example, introduces a framework to verify workflows, however, it is only applicable in a standardized scenario. Beyond that, van der Aalst et al. (2005) present process support strategies for unstructured processes in which the unstructured parts of the process are handled as individual cases.

In more detail, the literature of business process redesign contains numerous papers with the focus on the management of standardized processes and workflows. Reijers and Limanmansar (2005) derive a conceptual framework with the goal of best practices in business process design. In their paper, they focus on the mechanics of the process rather than on behavioural or change management aspects. They present a number of concrete redesign goals but their concept remains very general. Mansar and Reijers (2007) continue this research in their paper. They base their results on an empirical analysis of the top-ten best practices in business process redesign and the development of a framework to classify different approaches.

Redman (1995) switched the point of view from the model in general to specific redesign goals. In his analysis he points out that data quality is a competitive advantage. Within the scenario of AT&T, he describes the structure of a process to identify and eliminate deficient data quality. Moreover, Davenport et al. (2004) intensify this quality focus to guarantee appropriate decision making by process integration and data basis improvement. An addition to these general approaches and overviews are structured methods like Workflow Management Systems (WFMS). As van der Aalst et al. (2005) show, the applied method depends a lot on the characteristics of the specific domain. Consequently, they provide different approaches either for structured or unstructured processes. In the structured domains, they suggest the application of WFMS (van der Aalst, 2000; van der Aalst and Weske, 2001) based on Petri Nets. Moreover, van der Aalst and Weske (2001) present an approach to handle collaborative processes via integration in an inter-organizational context. They identify two characteristic processes: first the globally visible process and second the private subprocesses of each participant. Analogously to Davenport et al. (2004) their main goal is the perfect integration of all process parts into one main process.

The redesign approach described in our paper is related to this inter-organizational idea. Nevertheless, van der Aalst and Weske (2001) present a theoretical model and in this way miss to provide concrete support in realizing the new process. In addition to WFMS, van der Aalst et al. (2005) suppose the workflow management by knowledge workers who are supported by a system that presents all available informations. In doing so, the system supports the decisions made but does no autonomous decision making. Altogether, van der Aalst (2000) state that not all processes can be transformed into a standardized system. Nevertheless, they do not present a solution for processes that can be standardized in parts.

Beyond theses process structure focussed studies only few studies exist with additional dimensions. For instance, Feldman (2000) claims that procedural routines can only be handled correctly taking the human part into account that oftentimes causes unexpected changes in static routines. Pointing into the same direction, Seidel (2009) presents results of an exploratory study assessing the role of IS in the field of creativity intensive domains. Even more important is the technology focus of Sabherwal and Robey (1993) since they present a detailed study of IS implementation processes. With their resulting taxonomy of six different archetypical processes they provide valuable insights into the variety of implementation approaches. Yet, the link back from implementation to conceptual redesign, necessary for the application scenario of this thesis, is again missing. The same shortcoming is true for the structural metrics introduced to process redesign by Balasubramanian and Gupta (2005). Nevertheless, they claim a "formal yet user friendly" redesign approach and in that sense are quite close to the requirements formulated in Section 3.2.1. Finally, Kettinger et al. (1997) provide a valuable formalization and classification framework for process redesign approaches. Their proposed six stages form the structural back bone of the objectives-based redesign model of this work.

Summing up, most of the above presented papers work on the management of existing IS solutions rather than on their evolution through the integration of new processes. In addition, they focus either on completely structured or unstructured processes, or they miss to provide a concrete process redesign model. This incompleteness results in the necessity for a model with a higher flexibility regarding degree of structuredness of the affected processes and model presentation. The need for an appropriate redesign model is even more striking since an optimal implementation strategy incorporating management support and user involvement is crucial for the success of a newly implemented process (Schultz, 1992). Nevertheless, the literature presents criteria for an efficient process like redesign goals, which we utilize in our work (cp. Section 3.1.2).

2.2. Data and Process Quality

The terms data and process quality have been defined in numerous ways. However, research has brought up a limited number of core dimensions during the last years that are valid for academics and practitioners. Section 2.2.1 describes the quality dimensions that matter in the next chapters in detail and compares the different points of view in literature. Based upon that, Section 2.2.2 presents the set of suitable key performance indicators (KPIs) for the measurement of all quality dimensions. Yet, KPIs established throughout this work, for instance, the conformity with Benford's Law, are not part of this review.

2.2.1. Definitions in Literature

The overall goal of this work is quality assurance in financial planning, i.e. the quality of the data underlying risk management. For appropriate output data quality, process structure and efficiency are of utmost importance (Schultz, 1992). Consequently, although we talk about data quality in the following only, process characteristics are important, too. The term *data quality* contains numerous sub-types. Especially the different views of researchers and practitioners result in a wide range of definitions. Table 3.1 presents the explicit mapping between process characteristics and data quality dimensions.

According to Wang and Strong (1996) the most popular metric is data accuracy. Nevertheless, multiple dimensions beyond data accuracy exist, like for instance, timeliness and completeness. These dimension differ a lot in their popularity in recent literature and can be grouped differently. In this work we focus on the most popular dimensions and follow the structure of Batini et al. (2009), who extracted the following definitions for the four core dimensions based on a detailed literature research:

Completeness "The degree to which a given data collection includes data describing the corresponding set of real-world objects." In financial planning, a transaction between two subsidiaries always is present in the data of both subsidiaries– the same has to be true in the corresponding planning data.

Consistency "The consistency dimension refers to the violation of semantic rules defined over a set of data items." Going back to the above-described example, the data has not only to be present in both plans, but has to be, for instance, in the same currency and with the same volume.

Time related dimension– Timeliness Batini et al. (2009) refers to different dimensions, but we chose the very popular timeliness, that is, whether the data is out of date (Wand and Wang, 1996; Ballou and Pazer, 1985). Additional consistency and completeness validations along with new deadlines require reduced time to keep timeliness. For short, we write reduced timeliness instead throughout the work.

Accuracy Strongly depends on the application domain and can be distinguished between *semantic* and *syntactic*. *Semantic* is the comparison to the real-world values and is appropriate in our context where we are interested in the difference between planned and actual values.

As already mentioned, these four dimensions are very common, both for academics and practitioners as illustrated in Table 2.1. Sometimes, the definitions vary a bit, yet, the differences are caused by the specific domain. This modified application domain is also the reason for the sometimes differing dimension sample in the practitioner literature. Usually these variations lead to extensions of the dimension set. For instance, Knight and Burn (2005) present a list of more than ten dimensions, including security, reliability, accessibility, availability, among others. Some of these additional dimensions are present in further literature, too (Gardyn, 1997; Mandke and Nayar, 1997).

The only real exception regarding the core dimensions are Matsumura and Shouraboura (1996), who only take *accuracy* into account. Yet, as argued by multiple authors that is not enough since assessing data quality beyond accuracy is crucial (Wang and Strong, 1996; Batini et al., 2009). Overall, the above-described four data quality dimensions are the core dimensions present in practitioners and academics literature and consequently for this work.

2.2.2. Measurement

The evaluation presented in this work focusses on the four quality dimensions introduced in the previous section. As we do not aim at cost reduction but rather at decreased runtime and increased quality, costs are required to remain constant. Nevertheless, they are considered indirectly through the KPIs related to runtime. For instance, waiting time reflects waiting costs.

Buchsein et al. (2008) and Kütz (2011) describe a set of KPIs for a best practice process. This best practice process is taken from the IT infrastructure library[1] being the most popular IT management framework (Lahtela et al., 2010; Brenner, 2006), offering guidelines onto design, management, and support for service providers. Earlier work in the financial service industry proposes numerous indicators for performance

[1] http://www.itil-officialsite.com

		Timeliness	Completeness	Consistency	Accuracy
Academics	Wand and Wang (1996)	X	X	X	X
	Batini et al. (2009)	X	X	X	X
	Ballou and Pazer (1985)	X	X	X	X
	Knight and Burn (2005)	X	X	X	X
Practitioners	Cykana et al. (1996)	X	X	X	X
	Gardyn (1997)	X	X		X
	Mandke and Nayar (1997)	X	X	X	X
	Matsumura and Shouraboura (1996)				X

Table 2.1.: Mapping data quality dimensions to respective academics and practitioners literature.

measurement (Spremic et al., 2008), of which we chose a limited set. In order to validate and reduce the large set of KPIs proposed in the framework, we performed semi-structured expert interviews among knowledge workers at our industrial partner's site to ensure the practical relevance of all indicators. The results are presented in Table 2.2 being sub-divided into literature-based and expert interview-based KPIs with some indicators induced from both sides.

The execution time of the complete process is expressed through the indicators *Processing Time*, *Waiting Time*, and *Planning Time*. In an interaction process, for instance, between a holding and a subsidiary, *Waiting Time* and *Processing Time* are considered from the holding perspective: *Processing Time* is defined as the time the holding is active. Such an activity is, for instance, the validation of the subsidiary's planning data. *Waiting Time* is defined as the time the holding is passive, e.g. it waits for a

	Buchsein et al. (2008) Kütz (2011)	Expert interviews
Timeliness	Processing Time (PrT), Waiting Time (WaT), Planning Time (PlT)	80% Resolution Time (ReT)
Completeness	Number of Cycles (NoC)	Number of Performed Validations (NoPV), NoC
Consistency		
Accuracy	Absolute Percentage Error (APE)	

Table 2.2.: KPI's divided in the four data quality dimensions according to Section 2.2.1.

subsidiary's response. Finally, *Planning Time* is defined as the sum of *Processing* and *Waiting Time*. In addition to these straightforward KPIs extracted from literature, the expert interviews suggested an indicator that documents the workload development of knowledge workers both within subsidiaries and holding. The resulting *80% Resolution Time* represents the time interval between delivery deadline and the completion of 80% of the considered subsidiary's plan. Thereby, the 80% benchmark is a value derived within the expert interviews. Furthermore, Kütz (2011) proposes the *Number of Cycles*, that is the number of completed request/response cycles between holding and subsidiary (i.e. one email from holding and subsidiary), to reflect data quality in the sense of *completeness* and *consistency*.

The expert interviews revealed that data quality is likely to increase with the number of validations. Since each validation causes communication activities, an increased *Number of Cycles* is likely to highlight quality improvements. Based thereupon, we defined the *Number of Cycles* as an indicator to be increased or at least to be kept on a constant level. Beyond the documented number of iterations the *Quantity of Performed Validations* itself can be determined to provide additional information (for details about the performed validations refer to Section 3.5.4 and Martin and Blau (2010)). Table 2.2 provides an overview of all KPIs mapped on the respective quality dimension. Thereby, the above-described indicators measure the process-driven improvement in the three dimensions *timeliness*, *completeness*, and *consistency*. For the data-driven improvement we apply the straight forward indicator *Absolute Percentage Error*, calculated as the percentage deviation between planned and actual amounts. The detailed calculation description of all indicators will be given in Section 5.1 for the *APE* and in Section 3.5 for the remaining values.

Although *accuracy* and *APE* is the metric that is oftentimes of utmost interest in practise, further measures are required to indicate the data quality during the data generation process. The main reasons are unpredictable events that can cause a high plan-actual deviation although the planning process and the data quality were perfect. One first step towards this objectivity is the set of quality dimensions described in the previous section. The next step and one main challenge of the further work is to derive ex-ante quality indicator that evaluate data quality directly after generation and benchmark them against existing ones.

2.3. Data Analyses

One goal of this thesis is the quality assessment beyond the well known ex-post quality measure accuracy. To tackle the resulting challenges and to understand the performed evaluations, knowledge of general data mining procedures is required as they are described in the final related work Section 2.3.1. In addition to this general data exploration approaches, Section 2.3.2 introduces two special data analyses techniques. Beyond a detailed analyses of existing literature working on Benford's Law, the idea of weak planning efficiency, an approach mainly applied in macro-economic analyses, is explained.

2.3.1. Data Exploration

Similarly to the three stages introduced in Batini et al. (2009), the most challenging part in data analysing and data mining projects is the documentation and data understanding part. Only with a deep knowledge of data and process structures in mind, a successful decision support is possible. Thereby, data not only incorporates process structures, but also reflects human behaviour within a planning process. Consequently, one of the most important facts in data exploration projects is the imbalance between preparation and data surveying as it is illustrated in Table 2.3 (close to Pyle (1999)).

The preparation phase comprising of problem and solution exploration and implementation specification takes only 20% of the total time, however, it is of particular success importance. Moreover, data preparation takes 60% of the time and provides again high success importance. In contrast, surveying and modelling data takes 20% of the time and consequently is of lower importance to success. Zhang et al. (2003) support this crucial role of data preparation. The data base in large enterprises often contains parts with low-quality data, e.g. incomplete data sets. This assessment is also present in the field of data mining approaches.

A large body of research has been performed in last decades to identify the most efficient data surveying and modelling techniques. For instance, Fu (2011) presents an extensive overview of research on data mining in general including a structured representation of existing time series analyses. One of their major conclusions is the fundamental problem arising from the necessity for an uniform representation of the time series from domains with heterogeneous dimensions. E.g. nearly continuous time series for share values containing an almost infinite number of points in time (dimensions) have to be handled as well as time series comprising of five plan items as they are present in this work. The effort of handling this diversity led to an expanding diversity of surveying and modelling techniques throughout the last years. Besides dimensionality reduction, these techniques can be categorised into similarity search, clustering, visualization, and mining. To provide an overview of the aris-

<table>
<tr><th colspan="2"></th><th colspan="2">Time to complete</th></tr>
<tr><td>1.</td><td>Exploring the problem</td><td>10</td><td rowspan="3">20</td></tr>
<tr><td>2.</td><td>Exploring the solution</td><td>9</td></tr>
<tr><td>3.</td><td>Implementation specification</td><td>1</td></tr>
<tr><td rowspan="3">4. Data mining</td><td>a. Data preparation</td><td>60</td><td rowspan="3">80</td></tr>
<tr><td>b. Data surveying</td><td>15</td></tr>
<tr><td>c. Data modelling</td><td>5</td></tr>
</table>

Table 2.3.: Percentage importance and duration of each stage in a data exploration project based on Pyle (1999).

ing variety in the single part of time series clustering, Warren Liao (2005) provides a characterisation of previous research in dependence of underlying data, applied similarity measures, and application domains. In parallel, multiple approaches have been developed to reduce the dimensionality of input data. Keogh et al. (2001) introduce a performance optimized approach and benchmark it against three major dimensionality reduction methods. Meanwhile, Fu et al. (2008) present techniques specified for the application in financial data. On a more abstract level, many initiatives have been conducted to agree on a common base of techniques and common data input structures (Grossman et al., 2002).

One further well-investigated research field applicable to detect systematic errors in planning data are fraud detection metrics. Actually, systematic errors and fraud can lead to the same remarkable data structures, for instance, extraordinary repetitions of the same number. Numerous data mining techniques are applicable to detect such patterns. Phua et al. (2010) collected the publications of the last ten years in automated fraud detection literature to derive a clustering of fraudsters on an organizational and motivational level within an enterprise. Besides the identification of the most successful data mining methods in detecting fraud, the technical nature of data places again special needs on applicable data mining approaches. In addition to the excellent general literature overview by Phua et al. (2010), many comparative studies of existing and introductions of newly developed data mining procedures can be found. Bhattacharyya et al. (2011) compare the predictive quality and the efficiency of support vector machines, random forecasts, and logistic regression in the domain of financial credit card fraud. Fanning and Cogger (1998) introduce self-organizing Artificial Neural Networks to detect management fraud. In order to identify the characteristics of managerial fraud, they evaluate the predictable strength of twenty publicly available variables like, for instance, the existence of an audit committee. Going back to the application domain of rolling financial planning, the analyses of plan revisions can provide valuable knowledge. For instance, Yelland (2006) presents a comparative study with the scope of stable seasonal pattern models to refine the incorporation of weakly sales numbers in quarterly business plan updates.

Summing up, throughout all application domains heterogeneous data structures raise technical challenges. In addition, understanding the human factor during a forecasting process plays a key role in an improvement process (Bretschneider and Gorr, 1989). One metric assessing human and political planning behaviour based on data revisions from one delivery to the next is the so called weak forecast efficiency (Nordhaus, 1987). Simply the investigation of systematic up or down corrections and their improvement with tools removing 50% of the positive revisions holds improvement potential of nearly 10% (Fildes and Goodwin, 2008). Another set of methods to detect anomalies mostly independent of the application domain and data specific technical challenges are digital analyses based on Benford's Law first introduced by Nigrini and Mittermaier (1997). Like all previously described data analyses techniques, weak forecast efficiency and conformance with Benford's Law found their way into a multitude of different applications. However, none of them has been introduced to corporate financial planning data. One reason might be the complex data structure of planning, actual, and real-world data that results in tremendous challenges for analyses methods.

Nevertheless, a successful extraction of planning patterns and the improvement in financial planning data requires the extraction of extensive knowledge about data structures and their understanding. To cope with the arising challenges, we introduce the two metrics with low requirements for underlying data – weak forecast efficiency and conformance with Benford's Law – to financial planning data and analyse the influence of different dimensions in the next section.

2.3.2. Candidate Metrics and Benford's Law

One of the main contributions of this work is the introduction of Benford's Law to financial planning data. Consequently, we have to prove its general applicability in the domain of financial planning.

The digital phenomenon today known as Benford's Law or the significant digit law was initially discovered and described by Newcomb (1881). Benford (1938) found the first empirical evidence for it. The heterogeneous data underlying his studies ranged from numbers on newspaper covers to physical constants. Contrary to intuition, Benford spotted that the digits of these numbers are not uniformly distributed but rather follow a logarithmic distribution. The probability $P_i(j)$ for each digit j in position i of a number can be calculated as follows (for $i \in \{1, 2\}$):

$$P_1(j) = log\left(1 + \frac{1}{j}\right) \tag{2.1}$$

$$P_2(j) = \sum_{k=1}^{9} log\left(1 + \frac{1}{kj}\right) \tag{2.2}$$

Besides the striking distribution for $i \in \{1, 2\}$, Table 2.4 also depicts the probabilities of the third and forth position to illustrate the approximation towards an uniform distribution for higher digit positions (Nigrini and Mittermaier, 1997). For an intuitive explanation of the above-described phenomenon, please refer to Drake and

Nigrini (2000) and Durtschi et al. (2004). As shown by, for instance, Carslaw (1988) and Nigrini and Mittermaier (1997), Benford-style analyses can also be made for the combination of two digits.

Hill (1995) proved that Benford's Law follows a systematic statistical behaviour: since data distributions in nature are usually random samples taken from random distributions and joined afterwards, they converge to the logarithmic distribution as shown in Equations 2.1 and 2.2. Based on that, Nigrini (2000) derived three criteria to decide whether a data sample is likely to comply with Benford's Law:

1. The numbers should describe the relative sizes of similar phenomena,
2. The numbers should have no fixed upper and lower boundaries, and
3. The numbers should not be systematically created and assigned, as, for instance, ID-numbers.

For high quality financial planning data these criteria are generally fulfilled. To be of great value financial planning data must at any point in time include all relevant information. This information is highly heterogeneous and occurs randomly. Based thereupon, companies can calculate expected invoices and cash flows. Hence, the resulting financial planning numbers are random themselves as they are based on different data sources with different random distributions (Hill, 1995). With this ideal data generation process in mind, all numbers included in a set of high quality financial planning data are *1.* cash-related (same phenomenon), have *2.* no pre-fixed boundaries and *3.* are not created systematically. Furthermore, Pinkham (1961) showed the scale invariance of Benford's Law, e.g. heterogeneous currencies as present in financial planning tasks should not influence the data's conformity to the expected distribution. Importantly, the above-mentioned criteria are necessary but not sufficient. Therefore, statistical analyses of relevant and representative data are still indispensable to ultimately test whether financial planning numbers satisfy Benford's Law or not.

Digit j	$P_1(j)$	$P_2(j)$	$P_3(j)$	$P_4(j)$
0	n/a	11.968%	10.178%	10.018%
1	30.103%	11.389%	10.138%	10.014%
2	17.609%	10.882%	10.097%	10.010%
3	12.494%	10.433%	10.057%	10.006%
4	9.691%	10.031%	10.018%	10.002%
5	7.918%	9.668%	9.979%	9.998%
6	6.695%	9.337%	9.940%	9.994%
7	5.799%	9.035%	9.902%	9.990%
8	5.115%	8.757%	9.864%	8.986%
9	4.576%	8.500%	9.827%	9.982%

Table 2.4.: Digit distribution in first to forth position in "naturally occurring" numbers according to Benford's Law (Benford, 1938; Nigrini and Mittermaier, 1997).

Up to date, the significant digit law made its way into numerous different domains and fields of application. In the end, Benford (1938) initiated an entirely new field of analyses which can be roughly categorized into two groups. Besides papers that provide additional mathematical insights and theoretical evidence for Benford's Law, (for instance Hill (1995)), researchers have dealt with its application to different kinds of data sets. Moreover, academia has brought forth a respectable body of empirical work that presents applications of Benford's Law, mostly to detect anomalies and fraud in data. Nigrini and Mittermaier (1997) define such a digital analysis as "the analysis of digit and number patterns with the objective of detecting abnormal recurrences of digits, digit combinations and specific numbers".

The conformity of data to Benford's Law has been shown for a couple of domains: for instance, Diekmann (2007) examines the digit distribution in statistical regression coefficients published in scientific literature. However, most work has been done in the field of accounting data. Carslaw (1988) and Thomas (1989) proved that reported earning satisfy Benford's Law; Nigrini (1996) detected conformity to Benford's Law in tax payments. Due to the development of digital analyses and mathematical investigations, scholar's conclusions on the practical relevance of Benford's Law are quite different. While earlier work, with Raimi (1976) being named as a representative, calls the observations "a curious mathematical phenomenon", more recent publications are aware of the value of Benford's results: Hill (1998) argues that digital analyses have a big impact on daily accounting business.

Benford's Law's worth for detecting errors and systematic procedures started with its application to accounting purposes. Based on deviations from the expected Benford distribution, Carslaw (1988) found evidence for systematically rounded up numbers in reported earnings of companies in New Zealand. Thomas (1989) backed up this finding with data from American companies. In addition, he found a systematic rounding down behaviour for reported losses. Checking available tax data against Benford's Law, Nigrini (1996) detected systematic mistakes in tax payments. Based thereupon, Nigrini and Mittermaier (1997) developed and evaluated a standardized set of procedures to analyze huge data sets through different kinds of digit distributions. Krakar and Zgela (2009) apply digital analyses to foreign payments in banking transfers. During the last decade, the continuous development of digital analyses and data mining techniques in general has been accompanied by a continuous growth of available data in all business areas. As stated by Rezaee et al. (2002), this development requires new auditing structures and systems providing continuous auditing procedures. Suggesting continuous auditing based on digital analyses, Nigrini (2000) points to the same direction.

We pick up these arguments by laying the foundation for efficient auditing systems to ex-ante validate financial planning, e.g. forecast data. This foundation is reflected in the assessment of financial planning data: as mentioned above, to date it has not been proven whether numbers from this domain satisfy Benford's Law or not, though the necessary criteria are met. Based on a large set of representative empirical data we show that the digit distribution in financial planning data is conform to Benford's Law. In addition, we conduct more detailed analyses of clustered data in order to evidence the suitability of digital analyses for quality improvement measures integrated into automated decision support services as a part of business intelligence systems (cp. Chapter 4). Finally and most striking, we show the ability of Benford's Law to indicate data accuracy (cp. Chapter 5).

To elaborate additional characteristics and to extend the applicable procedures, we introduce the additional quality metric *weak planning data efficiency*. Later on, we correlate the two of them. Data efficiency can be measured when a forecast for one specific outcome i is done repeatedly using forecast revisions, calculated as the difference between two consecutive forecasts:

$$rev_{i,t} = forecast_{i,t} - forecast_{i,t-1}. \tag{2.3}$$

The best possible forecast incorporates all available information at time t. However, we cannot test if forecasts fulfil this requirement and if all information is included at any time. But what we can do anyway is to test whether the revision process follows predictable patterns or not. If we observe predictable patterns, it does indicate that the revisions are not fully rational and could be improved.

In the forecasting literature Nordhaus (1987) introduced the concept of weak-form forecast efficiency. Weak-form efficiency requires that forecast revisions and errors are uncorrelated with past forecast revisions and errors (otherwise, errors and revisions would be predictable). Expressed differently, revisions and errors should follow a random walk. Furthermore, he proposed the following OLS regression, to test for weak-form planning efficiency:

$$rev_{i,t} = \alpha * rev_{i,t-1} + \varepsilon. \tag{2.4}$$

If the estimate α is close to zero, revisions are not (linear) predictable by preceding revisions. If one of the estimates is significantly different from zero, the forecast revisions would not follow a random walk and one could assume a bias. Analogously to Benford's Law, weak- form forecast efficiency is a common concept applied to test the rationality of macro-economic forecasters with respect to different dimensions. Fildes and Stekler (2002) analyse the current state of the art in efficiency research compared to alternative time series models. Based on analyses of historical US and UK forecasts they detect shortcomings in the application of the weak efficiency approach and conclude with areas holding the largest potential for improvement. Schuh (2001) presents a detailed analyses of under- and over-prediction of multiple indicators such as unemployment and inflation to find out whether some forecasters generate a better forecast than others. In contrast, Batchelor (2007) elaborates cross-country differences on the level of forecast efficiency. To generalize the previous findings, Lux (2009) derives a framework to model a collective decision process reflecting social dynamics in stochastic processes. Finally, all authors find evidence for inefficiencies in macroeconomic forecast leaving space for improvement.

Part II.

Business Process Optimization

Chapter 3.

Corporate Financial Planning Redesign – A Case Study

The optimization of financial planning processes contains multiple steps, whereby the fundamental first step is the optimization of the underlying planning process structure. Since such processes are of high relevance for the daily business, it is reasonable to apply a redesign model that meets the special requirements in this application domain, for instance, flexibility. The redesign presentation in this chapter starts with a problem description in Section 3.1, including a motivating example and a formalized representation of the state of the art redesign goals in literature. Section 3.2 presents the detailed description of an appropriate redesign model, containing a design-science based derivation of the model and a qualitative evaluation through predefined requirements. The following three Sections 3.3 to 3.5 present the case study structured into the redesign model application in Section 3.3, a service-oriented implementation of the predefined services in an IS (3.4), and finally the results in Section 3.5 evaluating the effects through the implementation.

3.1. Problem Statement

Semi-structured processes in general as well as the financial data transmission and interaction process focused on in this thesis are often-times driven by historically grown organizational characteristics in multinational enterprises. Due to the variety of process characteristics, the process redesign and optimization in processes with numerous workflow patterns is highly complex (Seidel, 2009). To date, literature as presented in Section 2.2 has put forth a large body of models and frameworks that tackle the measurement and classification of routines in processes and, hence, enable the classification of semi-structured processes and process redesign in general. Yet, they are all of theoretical nature and hardly provide hands-on advice for flexible redesign necessary in practical applications (Martin et al., 2011). To address this research gap and to equip practitioners with a flexible redesign model, the following sections propose a objective-based process redesign model. It is an approach for semi-structured, non-standardized processes, which inherits both aspects from WFMS and case handling and integrates the idea of sequential patterns.

3.1.1. Motivating Example

In this subsection, we exemplary describe the above-mentioned challenges at concrete instantiation of a semi-structured process: the data transmission and interaction process as a part of the corporate financial planning within our industrial partner. It is depicted in Figure 3.1 as it looked like before redesign. Although the illustration looks structured, it represents a semi-structured process since it is an idealized and aggregated representation of underlying workflows. Actually, the behaviour of all human participants strongly depends on personal preferences. For the graphical illustration we use the Business Process Modelling Notation (BPMN), which has become the de facto standard in academic and practice communities for business process modelling (Recker, 2010; Wohed et al., 2006). Furthermore, BPMN meets our requirements in representing collaborative processes (White, 2004) between local legal entities and central management.

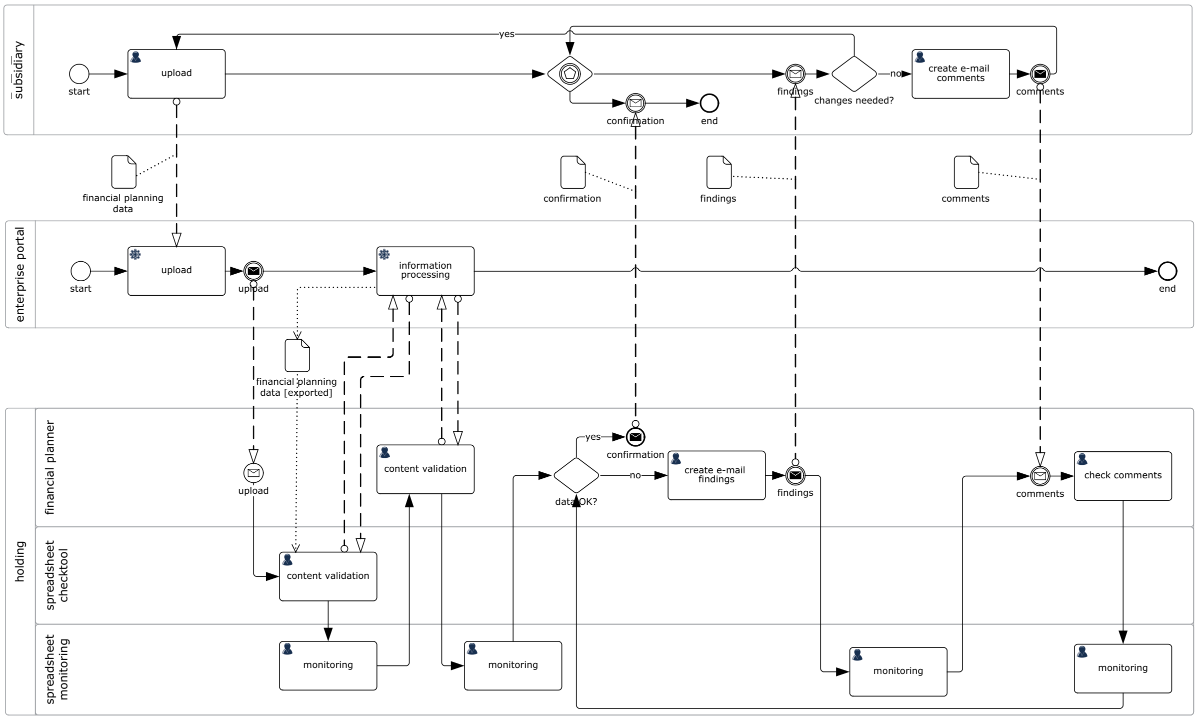

Figure 3.1.: Traditional data transmission and interaction process during financial planning within the industrial partner.

The process comprises of three pools representing local legal entity (subsidiary), gateway (enterprise portal) and central management (holding). Generally, the gateway could also be email communication in the most simple case or an IS in an advanced stage of redesign. The holding pool is again divided into three swim lanes. Activities in the lower two swim lanes are performed mainly with spreadsheet applications for data processing and monitoring. The upper swim lane represents manual process elements performed by knowledge workers in the holding company and is one source of unstructuredness: order and number of repetitions for all tasks depend on the respective knowledge worker.

The depicted process starts with the subsidiary sending financial planning data to the holding through the portal. This upload includes an automated validation regarding the structure of the delivered data. If successful, it initiates an upload notification in form of emails for both parties, that again initiates a detailed content validation on holding side. Meanwhile, the subsidiary's process has to wait for the result of the check performed by the holding company. The validation has to be carried out manually by a financial planner: *(i)* the financial plan is imported into the check-spreadsheet, *(ii)* the information is extracted from the portal and a monitoring-spreadsheet, and *(iii)* the validation results are documented in the corresponding monitoring spreadsheet and communicated to the subsidiary. In the following, the subsidiary has two options: to correct mistakes that caused the results or to enter comments if there are special issues reasonable for the results. Comments can be transmitted in an email; corrections in the financial plan data lead to a new upload and validation. This process part is another source for unstructuredness: kind of data generation and number of participants strongly depend on organizational structures within the subsidiary. Yet, if the data causes no results within the validation on the holding side and/or the comments are valid, the subsidiary gets informed by the holding and the process is finished.

Even in this simple example, standardization and automation as they are quoted in the following section hold strong potential for improvement: an automated implementation of the validation and monitoring tasks in the portal would eliminate waiting and processing times and reduce the number of participants. Manual tasks are often complex and very time consuming. They are difficult to communicate to colleagues and they hold a high potential for errors. Furthermore, an increased process integration would lead to an uniform data standard and provide organizational structures necessary for integrating the results of the data analyses in later chapters.

3.1.2. Redesign Goals

Although the redesign approaches described in Section 2.1 differ a lot in their application scenario, they offer multiple recurring criteria for an efficient process. In the context of a newly developed redesign approach these criteria can be applied as redesign goals to provide a theoretical backbone: Reijers and Limanmansar (2005) try to get rid of *(i)* unnecessary tasks, *(ii)* reduce contact and *(iii)* reduce waiting times.

Moreover, Redman (1995) presents solutions focused on *(iv)* task automation. In addition, data completeness in particular and data quality in general often depend on the process integration level. Therefore, van der Aalst and Weske (2001) as well as Davenport et al. (2004) claim the need for an increase of the *(v)* level of integration. Finally, Balasubramanian and Gupta (2005) present a structural metric for business processes containing most of the above-mentioned objectives. We denote the complete set of $m = 5$ objectives as $\mathcal{O} = \{O_i | i = 1, ..., m\}$. To expand the focus on data quality, Table 3.1 presents the link between the above-described five objectives and the four data dimensions relevant for this work (cp. Section 2.2 for a detailed data quality description). Therein, X indicates a positive influence of the objective on the respective quality dimension in a correct redesign application.

In this vein, increased *automation* can result in reduced *timeliness* and increased *completeness* and *consistency*, for instance, through the implementation of consistency and completeness validations as they are described in Martin and Blau (2010). The same relationship is true for *integration* eliminating system brakes and *reduce waiting times*. In contrast, the effects of *eliminate unnecessary tasks* and *reduce contact* are focussed on reduced time effort. Finally, all objectives are expected to increase data accuracy as the overall goal on the long run.

3.2. Objectives-Based Process Redesign Formalization

The redesign model presented in this section is a *design science artefact* and is part of the overall Business Process Redesign Framework derived in this chapter. This framework is based on the necessity shown in Section 3.1.1 along with the fundamental objectives in business process redesign (3.1.2). It comprises of the seven guidelines to be followed when pursuing a design science approach as introduced by Hevner et al. (2004). Consequently, the structure of this section is based on these design science guidelines. The first subsection presents the guideline “design as an artefact” in detail as the focus of this work is put on a model of redesigning semi-structured

	Timeliness	Completeness	Consistency	Accuracy
Automation	X	X	X	X
Integration	X	X	X	X
Reduce Waiting Time	X	X	X	X
Eliminate Unnecessary Tasks	X			X
Reduce Contact	X			X

Table 3.1.: Dependency between redesign objectives $\mathcal{O}$ and the data quality dimensions.

processes. Within this section, we apply the stages and activities of a business process redesign as presented and evaluated by Kettinger et al. (1997). The remaining design guidelines postulated in Hevner et al. (2004) complete the Business Process Redesign Framework in Section 3.2.2.

3.2.1. Redesign Requirements

The previous section introduced the necessity for an alternative redesign model and summarized the state of the art objectives in literature. Based upon this groundwork it is now possible to derive requirements posed to a redesign approach answering **RQ 1.1** (*How should a theoretically based process redesign model that combines standardization and flexibility be designed to assure practical relevance?*).

Consequently, this section provides a formalization to strengthen practical and theoretical relevance of the developed business process redesign model. We measure the achievement of these two goals based on the fulfilment of a set of $n = 6$ requirements $\mathcal{R} = \{R_i | i = 1, ..., n\}$. These requirements can be described as follows:

R1 *Objective conformity:* if possible within the constraints of the specific domain, the procedure must be able to realize all defined objectives $\mathcal{O}$.

R2 *Structured model:* the structure of the presented model should follow an accepted framework to support its research rigour.

R3 *Profound design methodology:* "the fundamental principles of design science research [...] are acquired in the building and application of an artefact" (Hevner et al. (2004)).

R4 *Flexibility:* the realization of objectives fractions $O \subset \mathcal{O}$ must be possible.

R5 *Simplicity of application:* a clear communication along with a structured representation of the model guarantee a simple application.

R6 *Applicable in Information System Design:* relevant redesign models must support the integration of existing processes into IS.

The requirements $R_1 - R_3$ in combination support research rigour and hence theoretical relevance of the redesign model. They reflect the state of the art in business process redesign (objectives in Section 3.1.2) and a profound methodology. In addition, the requirements $R_4 - R_6$ aim at flexibility and appropriateness for practical application as it is claimed in Section 3.1.1. In total, the requirements $\mathcal{R}$ fulfilment ensures an innovative redesign model. The derivation of such a redesign model in the following section is embedded into a design science framework along with the evaluation in the following chapter.

Hevner et al. denote design science as a problem solving process in which knowledge and understanding of a problem "and its solution are acquired in the building and application of an [IT] artefact" (Hevner et al., 2004). According to Hevner et al. (2004) and Walls et al. (1992), the definition of an IT artefact includes "not only instantiations [...] of the IT artefact but also the constructs, models, and methods applied in the development and use of information systems". Furthermore, Tsichritzis (1997) and Denning (1997) denote an IT artefact as innovations that define the idea, practices, technical capabilities, and products that are enabling the effective and efficient analysis, conceptualization, and utilization of information systems. In this vein, the redesign framework in this work contains an instantiation in Chapter 3 and the underlying model in the following section.

3.2.2. Design as an Artefact: The Redesign Model

The redesign model follows the stage-activity framework for business process reengineering as introduced by Kettinger et al. (1997). Their empirically derived work provides an enhancement of earlier fundamental work presented by Davenport (1993) and Grover et al. (1995) which includes a comprehensive survey of commonly used business process reengineering techniques and tools both from academia and business. The stage-activity framework for business process reengineering is composed of six stages of which our redesign model inherits five steps as detailedly shown in the remainder of this section. The evaluation stage was removed here since it matches the correspondent design guideline described in Section 3.2.2.

Stage 1 - Envision: each redesign project begins with the commitment and decision of the management. Redesign opportunities are discovered, suitable IT-related levers are identified and the targeted process is selected (Kettinger et al., 1997). In Section 3.1, we already introduced a motivating example for our redesign model's application. Analogously to this example, the redesign model is designed for *semi-structured processes*, which are non-deterministic sequences of activities: a semi-structured process is somewhere in between of ad-hoc and structured processes (Dustdar and Gall, 2003). Managing semi-structured processes requires a high level of flexibility, since they are not fully standardised, however, bring along a higher degree of structure than an ad-hoc process. The latter allows for the application of known activities, tools, and methodologies, yet requires a dedicated consideration of "fuzziness" (cf. Requirement R_2).

Stage 2 - Initiate: having identified and selected the field of application and the process to be changed, it is necessary to plan the redesign in detail and to define performance goals by analysing and determining the redesign requirements (Kettinger et al., 1997; Balasubramanian and Gupta, 2005). In the redesign model, the determination of performance goals (functional and non-functional) for the identified artefacts is defined by the set of objectives $\mathcal{O}$ as a structured representation of the general optimisation measures listed in Section 3.1.2. Hence, integrating the most fundamental general issues mentioned in literature leads to $m = 5$ objectives that need to be considered in (semi-structured) business process redesign.

Stage 3 - Diagnose: the initial state of the process including its sub-processes has to be documented prior to redesign (at time $t = 0$). We index the sequential redesign steps by $t \in \mathbb{N}$. Let $\mathcal{D}$ denote the *domain* of the process containing all process related information such as process attributes, resources, communication, roles, and IT (Kettinger et al., 1997). $\mathcal{D}$ is the only static documentation element since the domain cannot be changed by redesign steps (i.e. the domain sets the overall scope of the process). Based on $\mathcal{D}$, our redesign model identifies two basic concepts to document the process state at each time t: the constraints C_t of the domain $\mathcal{D}$ and the shortcomings S_t of the process. An example for a constraint is a limited automation degree that allows only for a few automated tasks during process runtime. C_0 denotes the initial set of limiting characteristics of the domain $\mathcal{D}$. All sets of constraints C_t with $t > 0$ are subsets of C_0. C_t impacts the process P_t at step t. These dependencies can be represented as mappings:

$$\mathcal{C} \;:\; \mathcal{D} \longrightarrow C_0, \tag{3.1}$$

$$\mathcal{P} \;:\; C_t \longrightarrow P_t,\; t \geq 0. \tag{3.2}$$

Deriving the initial set of shortcomings S_t includes, first, the domain-specific process P_t, and, second, the general set of objectives $\mathcal{O}$ (cf. Stage 2 - *Initiate*). The set of shortcomings S_t can be formalized as a mapping:

$$\mathcal{S} : (P_t, \mathcal{O}) \longrightarrow S_t. \tag{3.3}$$

In a nutshell, stage 3 is based on $\mathcal{D}$ and consists of the derivation of process specific shortcomings S_0 (the instantiations of the objectives $\mathcal{O}$ not fulfilled in the initial process P_0), and the constraints C_t. To exemplify the instantiation, we assume that there are three system brakes in P_0. In this case, S_0 contains 3 different shortcomings of the class $O_4 =$ *process integration*. In the following stage of our redesign model, we present an algorithm that deals with the documented shortcomings based on a stepwise constraint relaxation.

Stage 4 - Redesign: in stage 4 the actual redesign takes place. This stage of our redesign model is iterative and repeats along with the reconstruction stage 5. Each iteration is called a *redesign step* and the first step is indexed by $t = 1$ since $t = 0$ defines the status quo. Within each redesign step t we start by reducing and simplifying respectively the subset of constraints to $C_t \subseteq C_0$. We assume that some of the constraints C_0 can be deleted or at least formulated less restrictively (e.g. because of current technical developments we can automate some process parts which were not automated at $t = 0$). According to the Equations (3.2) and (3.3), the reduced constraint set C_t leads to a new process P_t and a new set of shortcomings S_t. Each redesign step t is successful, if $S_t \neq S_{t-1}$ holds. The redesign iteration will be stopped as soon as $C_t = C_{t-1}$ at a certain time t (i.e. the set of constraints cannot be reduced or simplified any more). We denote the index of the last executed redesign step by T. The algorithm including the exit condition is depicted in the following:

```
bool terminated = false;
int t = 0;
List<ConstraintSet> C = new List();
List<Process> P = new List();
List<ShortcomingSet> S = new List();
C(0) = getC(D));
while (!terminated)
  P.add(getP(C(t)));
  S.add(getS(P(t), O));
  C.add(C(t));
  C(t+1) = relaxC(t);
  if (C(t+1) == C(t));
    then terminated = true;
  t = t + 1;
end while
```

Executing this algorithm, we get an optimal process P_T with respect to the constraints C_T. The lists defined in rows (3) to (5) contain a documentation of the processed redesign steps, starting with the status quo in position 0. With k_t denoting the cardinality of C_t and l_t the cardinality of S_t it holds that $C_t = \left\{c_t^i | i = 1, ..., k_t)\right\}$ and $S_t = \left\{s_t^i | i = 1, ..., l_t)\right\}$.

Stage 5 - Reconstruct: the reconstruction consists of the realisation of the new process and its implementation in supporting IT-systems. As mentioned above, stage 5 forms the iteratively executed redesign step together with stage 4. However, only the implementing of a successful redesign, that is, the redesigned process has at least one shortcoming less ($S_t < S_{t-1}$), generates benefit. Consequently, non-successful redesign steps only contain stage 4.

3.2.3. Fulfilment of further Design Science Guidelines

After introducing *Design as an Artefact* as central design guideline to this work, the remaining design guidelines as proposed by Hevner et al. (2004) are summarized and mapped to our Business Process Redesign Framework in the following.

Problem relevance: the overall goal of design science research is not only to provide profound methodology, but also to develop technology-based solutions to relevant, that is, important business problems (Hevner et al., 2004). Continuously growing and changing multinational companies oftentimes struggle with heterogeneous degrees of standardization (Martin et al., 2011). Especially in case of redesigning business processes that have been historically grown over decades, the requirement of handling semi-structured processes is central. The crucial issue in practice and in theory is the integration of new processes into existing IS (Martin et al., 2011). In Section 3.1, we defined the challenges of such a procedure as requirements for the Business Process Redesign Framework. Still, academic literature does not provide a flexible tool, or framework, to handle such issues. The lack of standardization hampers the application of WFMS (van der Aalst et al., 2005), yet, proposed case handling approaches such as Löffeler et al. (1998) exhibit the major shortcoming of supporting the automation of tasks. Thus, the framework presented in this work can be rated both relevant in terms of business applicability and novelty.

Design Evaluation: Hevner et al. (2004) list different possible ways of performing an evaluation. In the previous section, we presented a general model for the redesign of semi-structured processes. In order to abstractly show that the redesign model fulfils the requirements stated in Section 3.2.1, Table 3.2 presents a brief explanation of the fulfilment and hence provides a qualitative evaluation. An instantiation of the redesign model as a design artefact adapted to a concrete use case is evaluated in Section 3.5.

Research Contributions: the research contribution is closely linked to the relevance of the redesign model. As above-mentioned, our model extends the present state of research by providing a defined procedure to tackle the ever-important issue of redesigning business processes in historically grown (IS) environments without limitations due to the processes' level of structuredness. The model is defined along a checklist that assures its theoretical and practical relevance (cf. requirements $\mathcal{R}$). As a contribution to academia, our redesign model yields *(i)* the problem representation in the diagnose step, *(ii)* the solution representation in the redesign step, and finally *(iii)* the design algorithm in the redesign step.

Research Rigour: design science is always a trade off between rigour and relevance (Hevner et al., 2004). As above-stated, our redesign model yields at a technique with high practical relevance. Nevertheless, a certain level of research rigour is achieved by thoroughly applying the well-established design science methodology by Hevner et al. (2004), with the stage-activity framework by Kettinger et al. (1997) as an embedded methodology to produce the design artefact.

Design as a Search Process: design science is also said to be an iterative approach that eventually satisfies the set requirements subject to the laws that constrain the problem environment (Hevner et al., 2004). Initially, the constraints as introduced in *Stage 3 - Diagnose* of Section 3.2.1 are restrictive. In several iteration steps, the

Requirement $R_i \in \mathcal{R}$	Fulfilment
R_1	Assuming that it is possible to get rid of all constraints ($C_3 = \{\ \}, t = T = 3$), the technique generates a redesigned process that fulfils all objectives $\mathcal{O}$ (Figure 3.4).
R_2	The redesign model is embedded into a Business Process Redesign Framework that follows universally accepted design science methodology by Hevner et al. (2004).
R_3	The redesign model itself fits into the stage activity framework presented by Kettinger et al. (1997).
R_4	Each redesign step realizes fractions of the objectives (cp. also Section 3.3.3) in abolishing shortcomings. The flexibility results from the ability of the redesign model to stop redesign at any time t and the flexible fraction size.
R_5	The structured knowledge representation ($\mathcal{O}$) along with the algorithmic description of the model support applicability.
R_6	The automation focus of the technique together with the representation of the domain in its constraints C_0 supports a stepwise service-oriented IS implementation (cp. also Section 3.4).

Table 3.2.: Evaluation of technique completeness based on the fulfilment of the requirements derived in Section 3.2.1.

constraints are relaxed, allowing for new automation steps. It is likely that some hard constraints given by the application domain remain in the final, redesigned process, thereby restricting the objectives. Yet, the redesign and reconstruct steps as above-described allow for a stepwise and flexible elimination of the shortcomings tailored to the environment of the problem and its constraints.

Communication of Research: the result of design science research shall be made available for both a technology-oriented and management-oriented audience (Hevner et al., 2004). We present sufficient details to allow an implementation of the framework in an appropriate application context (technology orientation Martin et al. (2011)) as well as the motivation why organizational resources should be committed to use the Business Process Redesign Framework in practice (management orientation Martin et al. (2011)).

3.3. Preliminary Stages

The case study presented in this section is the result of applying the redesign model from Section 3.2.1 to the semi-structured financial data integration process at our industrial partner. Consequently, it contains each of the six stages, whereby stage one (Envision) to four (Redesign) of the model are contained in this section. The detailed description of stages five (Reconstruct) and six (Evaluation) follow separately in Section 3.5.

3.3.1. Envision and Initiate

This section firstly defines the concrete instantiation of the process that is subject to redesign: the financial planning data transmission process as it has been described in detail in Section 3.1.1 and depicted in 3.1. The following sections develop multiple redesign steps based on this traditional version of the process.

Secondly, with the decision for one target process in mind it is necessary to define redesign goals and indicators measuring changes. These are mainly the general objectives $\mathcal{O}$ defining the non-functional redesign goals that we follow in our case study (cp. 3.1.2). In addition, the documentation of the redesign success in the next chapter for the three dimensions *timeliness*, *completeness* and *consistency* is based on the set of key performance indicators introduced in Section 2.2.2 and summarized in Table 2.2.

3.3.2. Diagnose

The Diagnose stage starts the redesign process in redesign step $t = 0$ with the documentation of the domain characteristics in the initial set of constraint C_0 (cp. Equation 3.1) and set of shortcoming S_0 (cp. Equation 3.3). For the reader's convenience and for a structured procedure representation, the description refers to the respective rows of the algorithm in Section 3.2.1.

Redesign step $t = 0$**:** the documentation of the status quo leads to the set of constraints listed in Table 3.3 (row 6). The constraints $c_0^1 - c_0^4$ reflect the low degree of IT support resulting in multiple manual tasks. c_0^5 arises within an multinational company having historically grown and heterogeneous processes. The instructions from the central management are mainly restricted to the process output with the consequence of autonomous black box processes within the subsidiaries. In total, the constraints result in the traditional process as it is depicted in Figure 3.1. Applying the objectives $\mathcal{O}$ to this process (row 9) results in $l_0 = 28$ shortcomings, again listed in Table 3.3. Each s_0^i is driven by exactly one objective o^i but each objective o^i can cause multiple shortcomings s_0^i. For instance, the objective o^4 *task automation* results in the shortcomings *manual result communication* (s_0^1), *manual data validation* (s_0^5), *manual comment validation* (s_0^9), *manual monitoring* (s_0^{13}), *manual send data* (s_0^{16}), and *manual planning data generation* (s_0^{23}). Furthermore, with a decreasing number of constraints, the number of shortcomings caused by each objective decreases simultaneously.

3.3.3. Redesign

Each of the further redesign steps require the relaxation of constraints. Consequently, the reduction of constraints is one central assumption of this redesign model. However, the application domain could cause constraints that cannot be relaxed. The stepwise reduction of the constraints can be observed in Table 3.3.

Redesign step $t = 1$**:** this redesign step starts by checking the termination condition (row 12). We assume that the data validation and the communication of the results can be automated (c_0^2 and c_0^3). Thus, our relaxation is successful and the redesign does not terminate. The remaining $k_1 = 3$ constraints result in the process depicted in Figure 3.2. Therein, the set of shortcomings S_0 is reduced. E.g., the remaining shortcomings caused by objective o^4 are: *manual comment validation* (s_1^1), *manual monitoring* (s_1^5), *manual send data* (s_1^8), and *manual planning data generation* (s_1^{13}). Due to this strong decrease of the shortcomings the constraints c_0^2 and c_0^3 were chosen to be relaxed. Hence, for a successful application of the redesign model, it is inevitable to

Redesign step (t)	Specifier (C_t/S_t)	Cardinality (k_t/l_t)	Components $\left(c_t^i \in C_t/s_t^i \in S_t\right)$
$t=0$	C_0	$k_0 = 5$	c_0^1 manual comment validation, c_0^2 manual data validation, c_0^3 manual communication, c_0^4 manual monitoring, c_0^5 autonomous subsidiary process.
	S_0	$l_0 = 28$	s_0^1 to s_0^4 manual result communication, s_0^5 to s_0^8 manual data validation, s_0^9 to s_0^{12} manual comment validation, s_0^{13} to s_0^{15} manual monitoring, s_0^{16} and s_0^{18} manual send data, s_0^{19} and s_0^{20} system brake data validation, s_0^{21} and s_0^{22} system brake monitoring, s_0^{23} and s_0^{26} manual planning data generation, s_0^{27} and s_0^{28} system brake data upload.
$t=1$	C_1	$k_1 = 3$	c_1^1 manual comment validation, c_1^2 manual monitoring, c_1^3 autonomous subsidiary process.
	S_1	$l_1 = 18$	s_1^1 to s_1^4 manual comment validation, s_1^5 to s_1^7 manual monitoring, s_1^8 and s_1^{10} manual send data, s_1^{11} and s_1^{12} system brake monitoring, s_1^{13} and s_1^{16} manual planing data generation, s_1^{17} and s_1^{18} system brake data upload.
$t=2$	C_2	$k_2 = 2$	c_2^1 manual comment validation, c_2^2 autonomous subsidiary process.
	S_2	$l_2 = 13$	s_2^1 to s_2^4 manual comment validation, s_2^5 and s_2^7 manual send data, s_2^8 and s_2^{11} manual planing data generation, s_2^{12} and s_2^{13} system brake data upload.
$t=T=3$	C_3	$k_3 = 0$	$\emptyset$
	S_3	$l_3 = 0$	$\emptyset$

Table 3.3.: Development of process characteristics C_t and S_t during algorithms completion time T.

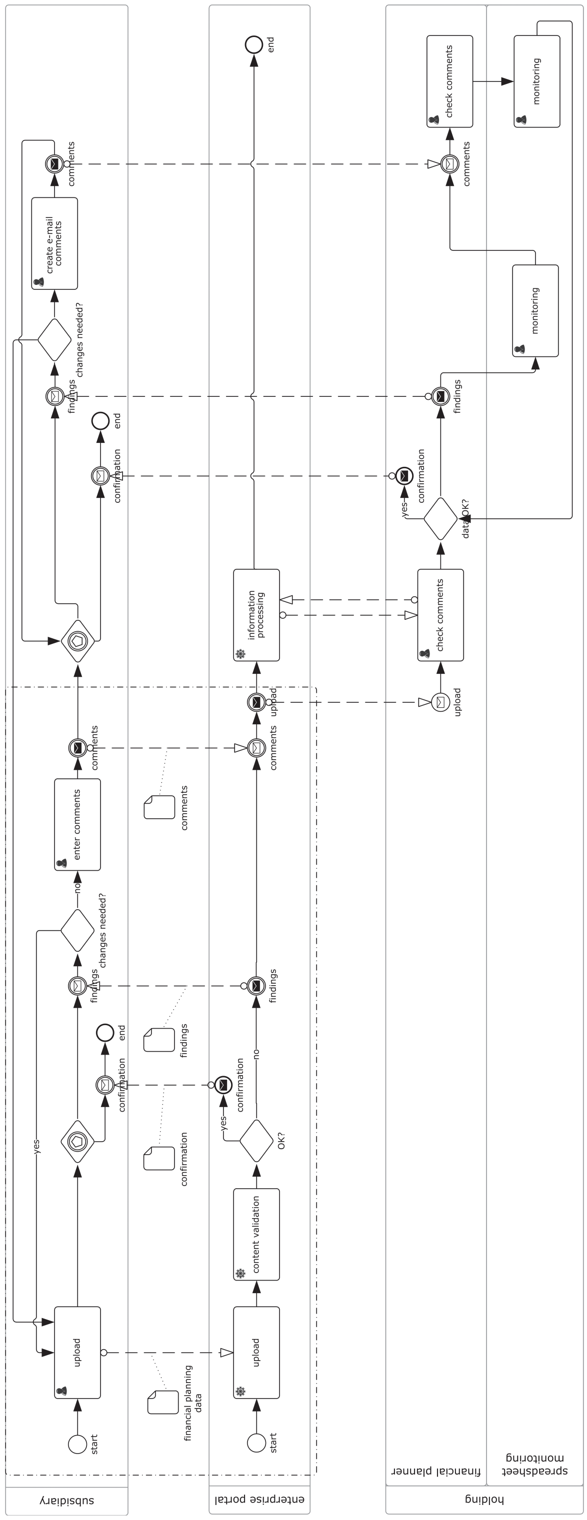

Figure 3.2.: Data transmission and interaction in financial planning after redesign step $t = 1$.

predict the benefits that arise from the relaxation of constraints. In doing so, even a semi-optimal process as it is depicted in Figure 3.2 will result in reduced effort. For instance, the automated validation in the holding process reduces costly knowledge worker resources. Moreover, manual data transfers between the enterprise portal and the spreadsheet application with a high potential error rate will be abolished. Altogether, the automation of the data validation promises to solve these shortcomings and therefore it is of high interest to automate such a task.

Redesign step $t = 2$**:** analogous to redesign step $t = 1$, the termination condition (row 12) is not hit. The further reduction of the constraints (cp. Table 3.3) allows for the automation of monitoring. This increases the process integration and abolishes another system brake. Considering again the shortcomings caused by objective o^4, only *manual comment validation* (s_2^1), *manual send data* (s_2^5), and *manual planning data generation* (s_2^{12}) are left. The resulting process is depicted in Figure 3.3. Therein, the data delivery process is completely automated including the data upload, the comment validation and the confirmations for both, the subsidiary and the holding. Furthermore, the enterprise portal also includes a continuous, detailed monitoring that offers real-time information about each aspect of the corporate financial planning process.

Redesign step $t = T = 3$**:** the remaining constraints $c_2^1 - c_2^3$ are very hard to relax due to the heterogeneous structure of the subsidiaries within our industrial partner. Nevertheless, for the sake of effectiveness and to demonstrate the completeness of the redesign model, we assume that they could be relaxed. In such an optimal scenario, the process would be entirely automated as it is depicted in Figure 3.4. Hence, the subsidiary would receive the results of the validations immediately after sending the data. All time-consuming tasks would be carried out in the enterprise portal and it would not be necessary to export data to another software application. Comparing the original process in Figure 3.1 with the optimized version in Figure 3.4, shows that all system breaks have vanished. Furthermore, at the end of this redesign step the process fulfils all objectives and it holds $C_3 = \emptyset$. Thus, no further constraint relaxation is possible and the algorithm terminates (row 13).

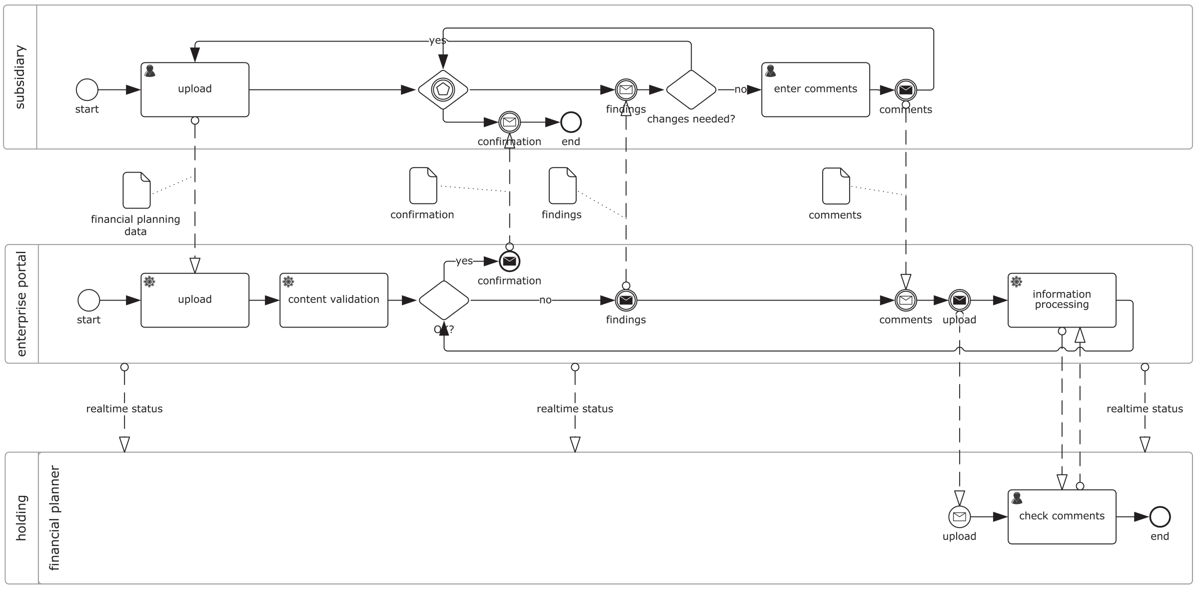

Figure 3.3.: Data transmission and interaction in financial planning after redesign step $t = 2$.

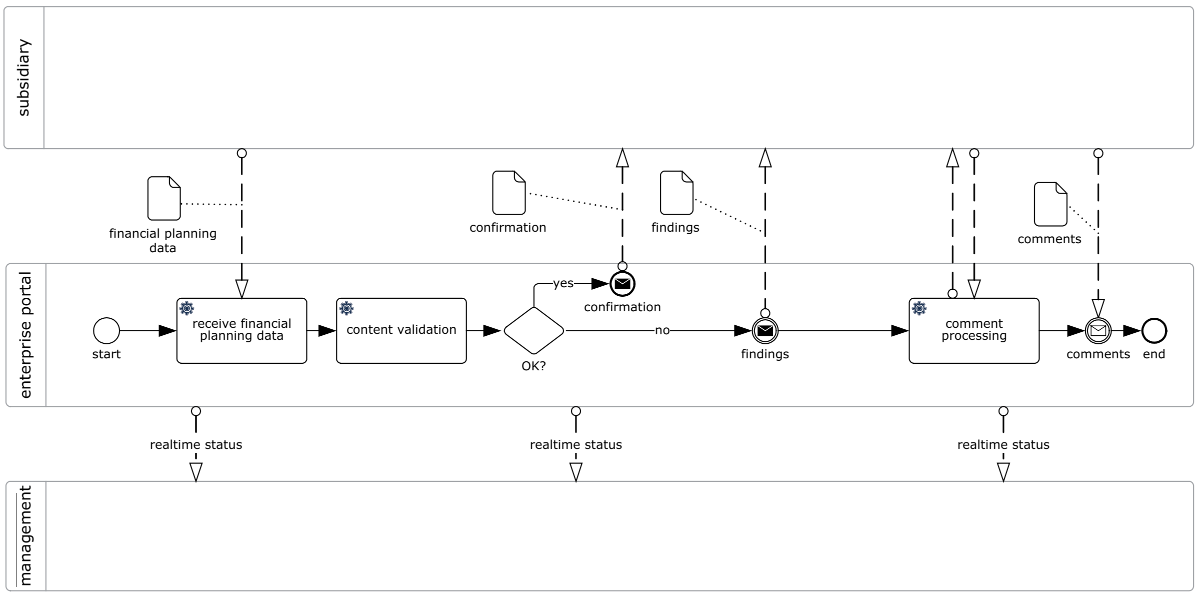

Figure 3.4.: Data transmission and interaction in financial planning after redesign step $t = T = 3$.

Altogether, the Redesign Model proposes at least 3 groups of tasks for automation and integration into the enterprise portal. As mentioned during the motivation of this thesis in Chapter 1, such enterprise portals usually are designed as IS. Consequently, this separation into granular groups of tasks strongly supports their service-oriented implementation in an IS as it will be presented in Section 3.4. The implementation of at least one of the presented services instantiates the respective reconstruct step. The final evaluation Section 3.5 in this chapter presents the results of the current reconstruct stage with the implementation of an *Upload and Validation Service*.

3.4. Technical Implementation & Redesign Status

The application of our redesign model for the given financial planning process results in different redesign steps as described in Section 3.3. The realization of these redesign steps requires the implementation of six generalised de-coupled service units, which are shown in Figure 3.5. These services can briefly be explained as follows:

A **Upload and Validation Service**, which serves two purposes: firstly, it enables the upload of financial reports (preformatted Microsoft Excel sheets) to the holding's report management system. Secondly, it facilitates the interactive validation of financial reports on-demand by a subsidiary. Hence subsidiaries can validate their reports before supplying them to the holding. Once a subsidiary has a valid report they may then upload it to the holding's report management system. However, as financial reports are typically complicated and whilst many aspects can be validated in an automated manner, there are often specific details that require additional explanation. A common example is a positive tax deduction, which could be due to a tax rebate. Therefore, many subsidiaries also upload reports that fail the validation process for one or more reasons. Such circumstances require additional meta-information (comments) to be provided by the subsidiary.

The **Comment Management Service**, therefore provides exactly this capability. This may be a set of standard reasons, or when there is no default answer set, free text is supplied by the subsidiary in the form of a comment.

The **Rule Engine**, is fairly self explanatory. A database stores the set of validation rules for the reports, and defines which comment sets relate to specific validation errors. The rule engine then enacts the rule set to performance the validation process, and afterwards, where relevant, determines which comment sets are of interest for the identified validation errors.

The **Management Service** acts as the access point for the holding's financial management team, and provides the capability for the holding to administrate financial reports and collect basic data, such as total cash flow and other financial key performance indicators. The administration of financial reports, in this context, means the ability of accounting staff to access reports, validation results and associated comments.

The **Monitoring Services** provide the ability for the holding to track the interaction and progress of subsidiaries within the financial planning service. Specifically, they enable the tracking of comment resolution, such that the holding's financial planners can follow a comment exchange, which may result in a revalidation process, and steer a subsidiary towards a successfully validated and submitted report on time. The *Monitoring Services* not only monitor the progress through each step of the service chain, but also provide real usage data that we can use to evaluate the effectiveness of process redesign, and allows the holding to identify potential bottlenecks in the process as a whole. It provides information such as: the number of validation iterations, time taken to perform validation, number of comment-related exchanges with the holding's financial planners etc.

The **Portal** is the Web-based central access point to the system. It is not a by product of the redesign process, but rather has been extended to enable access to the **Business Services** level (Figure 3.5) of the services defined in this section.

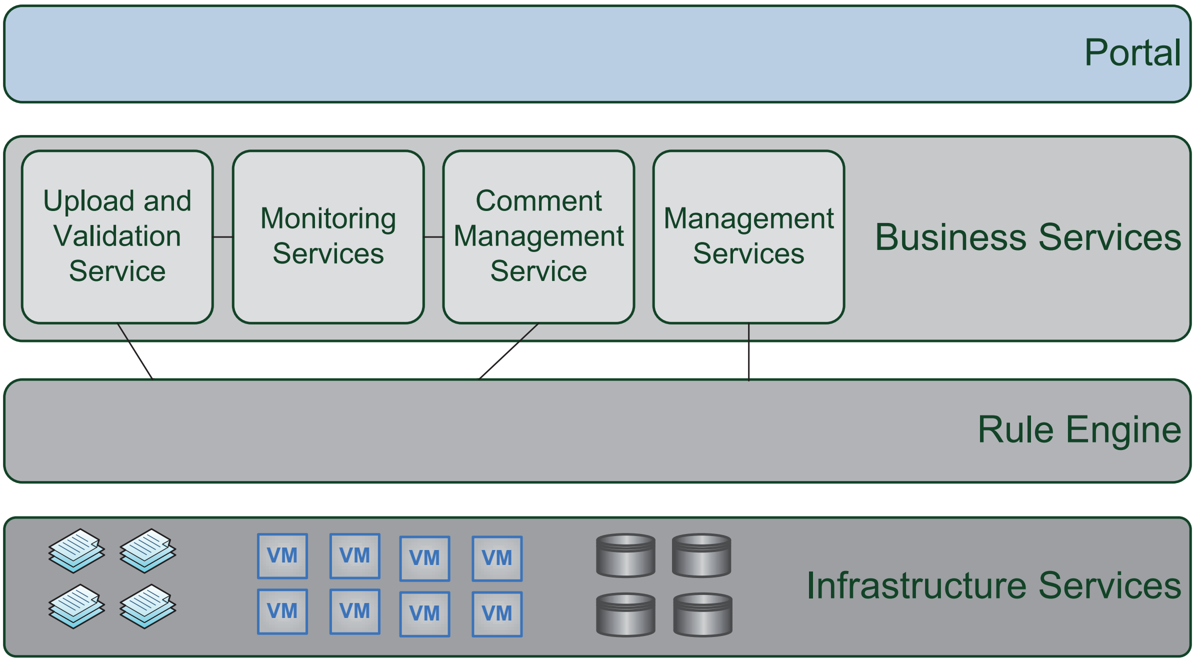

Figure 3.5.: Services Stack/Organisation in financial planning.

To support the above-described services, a set of infrastructure services is needed. In the context of our industrial parter these services are an in-house dedicated server farm. The results of the redesign process is, however, not sensitive to the backend infrastructure, which means that the holding could later move to a private Cloud setup if appropriate. Due to the sensitivity of the data, a public Cloud would likely be inappropriate. Ultimately, the infrastructure hosts each service unit, and particularly in the case of the validation and management services provides the needed computational resources for analytical tasks. These tasks, currently performed as batch jobs, are also candidates for parallel infrastructures as they are inherently parallel. This is specifically important if either of these services become performance bottlenecks in future reporting rounds, as the functionality that they provide can easily be scaled out in a parallel, distributed or cloud computing manner.

By combining each of the services into a single service chain, it is possible to completely replicate the original process model, but with a much higher level of automation and scalability. Note, however, that it is oftentimes not possible or not desired to introduce a completely new system into the business environment at once. This aversion is due to inherent monetary risks of rapidly changing the implementation of a business-critical process completely. Moreover, introducing monolithic systems can lead to dissatisfaction of staff. Therefore, the different proposed service modules are taken into the productive system gradually over time so that the performance and reliability of each module can be ratified and demonstrated.

Presently, the rolled out implementation is the *Upload and Validation Service*, whose performance is evaluated in Section 3.5. Nevertheless, simply taking this single service into the productive system fosters the further analyses in Chapter 3 to 5. Firstly, automated and hence complete validations assure a certain level of data quality. This development is strengthened by the additional manual validations that can be performed as one result of time savings through standardisation (cp. Section 3.5.4). Secondly, the implementation of an *Upload and Validation Service* establishes the organisational and structural prerequisites for the integration of a decision support service. Such a service would provide recommendations to subsidiary and holding based on information gained through analyses of historically available planning and actual data.

3.5. Reconstruct and Evaluation

In this section, the research questions **RQ 1.2** and **RQ 1.3** are evaluated via empirical data accumulated during the real-world application of the redesigned financial planning process at our industrial partner's site. Section 3.5.1 presents the underlying data sample, followed by a brief overview of the applied methodology in Section 3.5.2. Section 3.5.3 introduces the hypotheses as the back bone of the core evaluation in Section 3.5 that concludes with a brief interpretation of the evaluation results and the implications to be drawn.

3.5.1. Empirical Data Description

The data that underlies the evaluation includes seven data deliveries, starting in June 2009. The data has been generated in regularly recurring time periods as shown in Table 3.4. Therein, the number of delivering subsidiaries ranges between 99 and 113, owed to mergers and acquisitions that took place during the evaluated time periods. 89 subsidiaries constantly delivered data in all seven periods, which serves as the data basis for the following evaluation. Furthermore, the data sample comprises of two data delivery eras: the pre-redesign phase, which includes four deliveries from June 2009 (06/09) to November 2009, and the post-redesign phase, including the three deliveries from June 2010 to November 2010. The data delivered in March 2010 was distorted due to redesign implementation activities that took place during this delivery period. Therefore, this data set was excluded from the evaluation. It is important to note that our industrial partner has not only incorporated the redesign approach for the conceptual reorganization of the planning process, but has also integrated the services into its daily business immediately after their implementation (three pre-redesign deliveries are included in the evaluation). The data set per period includes the entire email communication between the holding and all subsidiaries and provides a detailed documentation of the validation and communication process. The email communication contains both manually and automatically generated messages. Automated notifications include information about the upload status and the validation results, while manual messages query, for instance, further planning data explanations.

The emails are classified into *Email Sender*, *Email Receiver*, *Email Subject*, and *Email Date*. That way, we can distinguish whether the email is sent automatically or not, whether it is sent by the holding or the subsidiary, and when it is sent. In this vein, an email sent by the holding to a subsidiary marks the switch from Processing Time to Waiting Time. Automated notifications are treated as generated on holding side, therefore, they also belong to this category. An email sent by a subsidiary marks the opposite switch, accordingly.

Delivery	06/09	09/09	11/09	03/10	06/10	09/10	11/10	Overlap
# Subsidiaries	99	100	106	104	113	113	113	89

Table 3.4.: Number of subsidiaries per data delivery during the evaluation with an overlap of *89* subsidiaries constantly delivering data from *06/2009* to *11/2010*.

Since the structure of the data sample is essential for the correct choice of statistical analyses, a careful examination of the data is required. According to the Shapiro-Wilk test for normality in SPSS 19, the distribution of the Planning Time is significantly non-normal (pre-redesign phase, $W(267) = 0.72$, $p < .001$; post-redesign phase, $W(267) = 0.68$, $p < .001$) and hence non-parametric interferential analyses have to be conducted (cp. next section).

3.5.2. Research Design

The evaluation presented in the following is founded on empirical data from the real-world application of corporate financial planning redesign with an evaluation period of nearly two years. Yet, we are aware that multiple effects that are not related to our redesign may cause positive and negative developments over such a long time period. The following measures aim at minimizing their influence.

Clustering: in the context of a multinational enterprise, our expert interviews have revealed that a time reduction of one or two working days has only a small business impact. To increase practical relevance, we performed a clustering approach and changed the data structure from working day intervals to working weeks.

In more detail, the original data (measured in working days) has been transformed into the corresponding number of working weeks after deadline with a cap at the end of the fifth week to smooth extreme outliers. Since we have only a very little number of extreme outliers (more than five working weeks), it is reasonable to include them in the fifth working week to reduce standard deviation in the empirical data. The transformation is defined by the following function:

$$t : R_+ \longrightarrow \{1,2,3,4,5\},\ t(x) = \begin{cases} \lfloor \frac{x}{5} \rfloor + 1, & \text{if } 0 \leq x < 20 \\ 5, & \text{else} \end{cases},$$

where $\lfloor \bullet \rfloor$ denotes the floor function. Based upon the resulting values $v \in [1, 5]$ we can calculate the average number of working weeks (compared to working days). However, the clustering can only be performed for KPIs measured in working days. It is not applicable to the Number of Cycles. Moreover, for 80%-Resolution Rate, which is a strongly aggregated value anyway, clustering is unlikely to generate additional insights. As we focus on working weeks where possible, we perform an inferential analysis on the clustered values only.

Comparability: planning data generation depends on multiple inputs, some of which include seasonal effect. At our industrial partner, for instance, new controlling numbers for the following year are available in November which are included during the forecast generation. Therefore, the information available for planning is not constant for all periods which requires a differentiation of the planning data by its delivery period. To avoid any mistakes caused by these seasonal effects, we proceed twofold: *(i)* we compare values of the same month (in 2009 and 2010), and *(ii)* we increase the comparability through the comparison of aggregated values per year.

Non-parametric analyses: for the inferential analysis we conduct the non-parametric Wilcoxon signed-rank test as the compared two data samples always include the same subsidiaries and the data sample distribution for all phases are non-normal (cp. Section 3.5.1). Based upon this inferential analysis we examine the deviations between the pre-redesign values in 2009 and the post-redesign values in 2010. According to Field (2009), we always add the 1-tailed level of significance *p* and the test statistic *T* (denoting the smaller value of the two rank sums) to the reported absolute value.

3.5.3. Hypotheses

In Section 3.2 we derived a theoretically based redesign model that is appropriate to fulfil the requirements listed in Section 3.2.1. To extend this qualitative evaluation, we now present a quantitative evaluation based on the implementation of the above-described redesigned process at our industrial partner. This evaluation aims at proving the improvement of three of our four quality dimensions – *timeliness*, *completeness*, and *consistency* – based on the key performance indicators described in Section 2.2.2.

In order to find evidence for **RQ 1.2** (*Does the objective-based process redesign increase the data quality dimension timeliness in practice?*) we investigate four different time dimensions in the data delivery process that are expressed in the following hypotheses (**H**):

H 1.2.1 – INNOVATIVE BUSINESS PROCESS REDESIGN –
Objectives-based redesign decreases processing time.

H 1.2.2 – INNOVATIVE BUSINESS PROCESS REDESIGN –
Objectives-based redesign decreases waiting time.

H 1.2.3 – INNOVATIVE BUSINESS PROCESS REDESIGN –
Objectives-based redesign decreases planning time.

H 1.2.4 – INNOVATIVE BUSINESS PROCESS REDESIGN –
Objectives-based redesign decreases 80% resolution time.

With the evaluation of these hypotheses we can address different dimensions of *timeliness* in the process: *(i)* the time effort on holding side through **H 1.2.1**, *(ii)* the time effort on subsidiary side with **H 1.2.2**, and *(iii)* the overall time effort. The overall effort is separately evaluate for time (**H 1.2.3**) and workload (**H 1.2.4**), which offers us a more detailed perspective on the changes in the process and strengthens our findings.

In a next step, the goal of the evaluation is to address two further dimensions in **RQ 1.3** (*Does the objective-based process redesign increase the data quality dimensions completeness and consistency in practice?*). In contrast to **RQ 1.2** it is much harder to derive appropriate indicators for *completeness* and *consistency*. Nevertheless, it is possible to find indirect evidence through the following hypotheses:

H 1.3.1 – INNOVATIVE BUSINESS PROCESS REDESIGN –
Objectives-based redesign increases number of cycles.

H 1.3.2 – INNOVATIVE BUSINESS PROCESS REDESIGN –
Objectives-based redesign decreases the number of performed validations.

While **H 1.3.1** can be evaluated on the same data foundation as the previous hypotheses, the investigation of **H 1.3.2** requires additional knowledge about the structure of the newly designed process and the content of the communication. Through our close cooperation with our industrial partner, we are able to gain both.

3.5.4. Results

This section presents the detailed evaluation results of the productive use of the *Upload and Validation Service* at our industrial partner using the available data and methodology described in Sections 3.5.1 and 3.5.2 to verify the hypotheses raised in the previous section.

Table 3.5 shows the results for the average Number of Cycles (NoC) and the average Processing Time (PrT). The latter is listed on a working day basis and on a working week basis (clustered). For each delivery period, Table 3.5 shows the KPI's values of 2009 and 2010 along with the relative deviation $\triangle(09,10)$. In addition, the overall line prints out the aggregated results over all planning periods per year, representing the average over 267 subsidiaries.

KPI/ Period	Average NoC			Average PrT			Average PrT -Clustered-		
	Value 2009	Value 2010	△ (09/10)	Value 2009	Value 2010	△ (09/10)	Value 2009	Value 2010	△ (09/10)
Jun	1.90	1.89	-1%	3.68	1.91	-48%	1.51	1.21	$-19\%^{**}$
Sep	1.48	1.72	$+16\%$	2.82	1.51	-47%	1.40	1.16	$-18\%^{*}$
Nov	1.66	1.91	$+15\%^{*}$	3.03	2.01	-34%	1.42	1.24	$-13\%^{*}$
Overall	1.68	1.84	$+9\%^{*}$	3.18	1.81	-43%	1.44	1.20	$-17\%^{***}$

Table 3.5.: Evaluation results: absolute KPI values (Average Number of Cycles, Average Processing Time unclustered and clustered) for 2009 and 2010 along with the relative deviation and the 1-tailed significance level ($p^{*} < .05$; $p^{**} < .01$; and $p^{***} < .001$).

Starting with **H 1.3.1** and the average NoC, Table 3.5 indicates no deviation in June, however, clear increases of NoC in September (16%), November (15%) and in the overall numbers (9%). NoC in November (Overall) is significantly higher in 2010 than in 2009, with $T = 497.50$, $p < .05$ ($T = 3678.50$, $p < .05$). For the average PrT (working days), the improvement through redesign varies from 34% in November to 48% in June. For the clustered average PrT, the improvement ranges between 13% and 19%. The PrT is significantly lower in 2010 than in 2009 for all four observations, $T = 73.50$ and $T = 192$, $p < .05$ in September and November, $T = 119$, $p < .01$ in June and even $T = 1079.50$, $p < .001$ in Overall.

Table 3.6 is structured analogously to Table 3.5 and provides the foundation for the investigations of **H 1.2.1** to **H 1.2.4**. The Average Waiting Time (WaT), the average Planning Time (PlT) and the 80% Resolution Time (ReT) clearly decrease for all periods in 2010 compared to 2009. For WaT, the reduction ranges between 4% in September and 11% in November. The relatively small reduction in September is likely to be caused by the small absolute value (1.44 and 1.38 working weeks). Moreover, the Overall WaT 2010 is significantly lower than in 2009, $T = 1723$, $p < .05$. The average PlT improves from 12% to 19% after the redesign. Again, Overall PlT 2010 is significantly lower than Overall PlT 2009, $T = 365.50$, $p < .05$ in November, $T = 317$, $p < .01$ in June and even $T = 2905$, $p < .001$ in Overall. Finally, ReT indicates a strong workload reduction. In 2009, ReT varied from 14 to 18 working days. In 2010, the highest

KPI/ Period	Average WaT -Clustered-			Average PlT -Clustered-			80% ReT		
	Value 2009	Value 2010	△ (09/10)	Value 2009	Value 2010	△ (09/10)	Value 2009	Value 2010	△ (09/10)
Jun	1.73	1.58	-10%	2.38	1.92	$-19\%^{**}$	18	13	-28%
Sep	1.44	1.38	-4%	1.91	1.67	-12%	15	11	-27%
Nov	1.56	1.39	-11%	2.00	1.72	$-14\%^{*}$	14	9	-36%
Overall	1.58	1.45	$-8\%^{*}$	2.10	1.77	$-16\%^{***}$	16	11	-30%

Table 3.6.: Evaluation results: absolute KPI values (Average Number of Cycles, Average Processing Time unclustered and clustered) for 2009 and 2010 along with the relative deviation and the 1-tailed significance level ($p^{*} < .05$; $p^{**} < .01$; and $p^{***} < .001$).

value is observed in June with 13 working days. With respect to the Overall ReT decreasing from 16 to 11 working days, we observe a reduction of one entire working week (which equals 30%).

The further discussion refers to the results in Figure 3.6 and Figure 3.7 that illustrate the changes in PrT and WaT for the first delivery after the *Upload and Validation Service* was entirely rolled out (June 2010). They depict the working hours per week for the first five weeks after the delivery deadline. According to the original values printed out in Table 3.5, the PrT in June 2010 decreases by 48% compared to the values measured in June 2009. This development is clearly reflected in Figure 3.6. In addition, Figure 3.7 illustrates the reduced WaT in June 2010 which decreases by 10% compared to June 2009. Figure 3.7 shows a second development which is certainly worth discussing: the WaT workload's balance point shifts to the left, from approximately 2.7 in 2009 to approximately 2.1 in 2010. This shift towards the delivery deadline is a direct consequence of the reduced Processing Time. The subsidiaries receive the results of the holding validations earlier than before and, hence, can start to work on their response earlier. The Planning Time as the sum of PrT and WaT is consequently affected by both above-described effects. Since the shift of WaT towards the delivery deadline suggests that the validation process itself relocates, the ReT is considered as the KPI to express a workload reduction or increase. Consequently, Figure 3.7 not only illustrates the improvement in WaT but also visualizes the before-mentioned re-

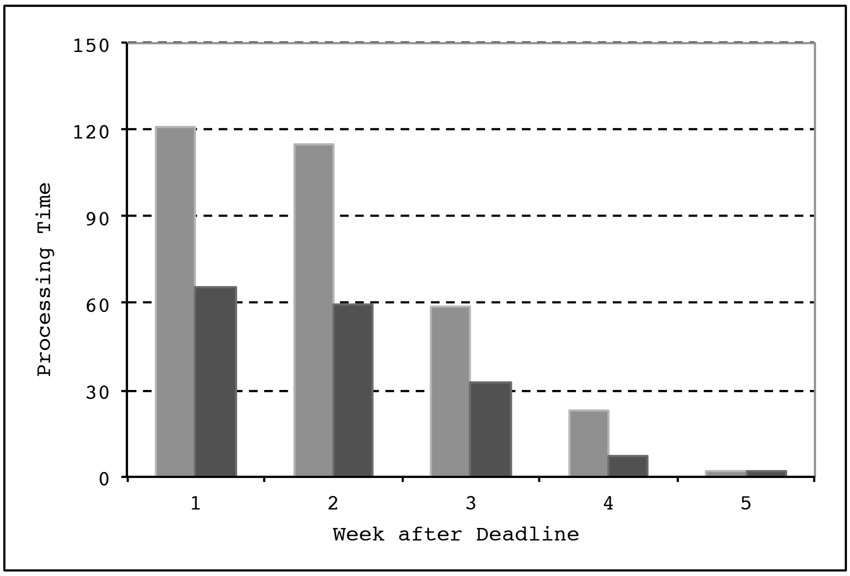

Figure 3.6.: Processing Time in hours per week after delivery deadline. Comparison between June 2009 (light grey) and June 2010 (dark grey).

duction of nearly one working week for all considered deliveries between 2009 and 2010: workload increases in the first two working weeks and reduces in the remaining three.

At large, all time-related KPIs (WaT, PrT, PlT) are significantly lower in 2010 (overall) than in 2009 ($T = 1723$, $p < .05$ for WaT, $T = 1079.50$, $p < 0.001$ for PrT, $T = 2905$, $p < .001$ for PlT). This reduction comes along with a one working week decrease in ReT. This improvement of all time-related KPIs is intensified when we look at NoC in the same period of time. NoC, and hence the communication activity between subsidiaries and holding, increases significantly in 2010 (overall) compared to 2009 (overall), $T = 267$, $p < .05$.

In addition, we are not only able to prove a quantitative but also a qualitative increase in communication: the automated data validations include both checks of *intra-subsidiary planning* and *inter-subsidiary planning* that were conducted manually prior to the redesign. The former denotes validations based on the planning data within a single subsidiary and related to only one single planning period. For instance, comparison of invoices and payments for a specific time horizon. The lat-

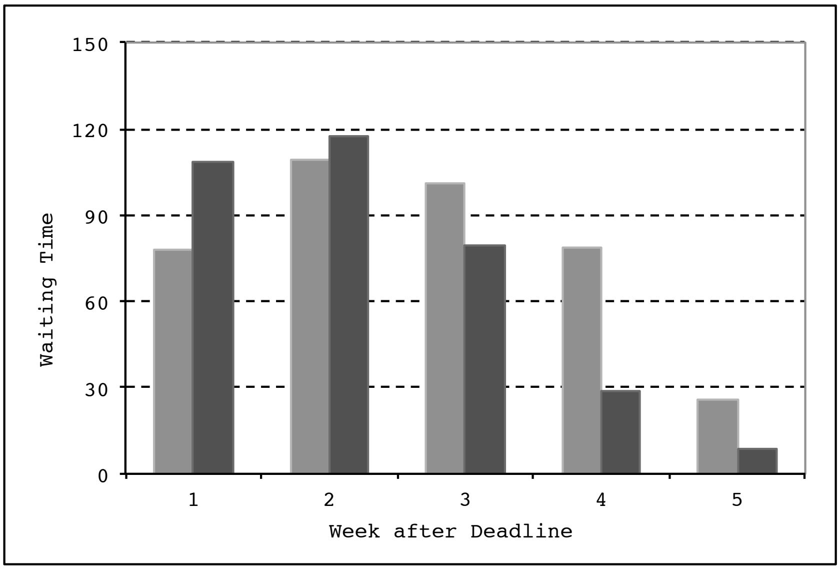

Figure 3.7.: Waiting Time in hours per week after delivery deadline. Comparison between June 2009 (light grey) and June 2010 (dark grey).

ter defines validations that take up data delivered by at least two subsidiaries, yet still relates to one single period. That is, for example, the consolidation of invoices issued and received between two entities. For a detailed description of such *completeness* and *consistency* validations please refer to Martin and Blau (2010). Along with the automated checks, the respective notifications of the validations have also been automated. Additionally, new and more sophisticated verifications of data deliveries over several planning periods (*inter-subsidiary/inter-period planning*) to further increase data quality, have been added. Table 3.7 compares the carried out validations pre- and post redesign in dependence of the entity size, whereby **–** indicates validation for no entity, **-** for few entities, **+** for most entities, and **++** for all entities. As easily can be seen, much more entity data gets validated regarding *completeness* and *consistency*.

Summing up, we have introduced a redesign model appropriate to formalize corporate financial planning redesign and have shown the strong data quality potential of such redesign through implementation at our industrial partner. Under the conditions of a real world application (as formulated in Section 3.5.2) we found indication for confirming all four hypotheses **H 1.2.1** to **H 1.2.4** regarding the data *timeliness* and

Validation type	Pre-redesign		Post-redesign	
	Small entities	Large entities	Small entities	Large entities
Intra-planning	++	++	++	++
Inter-planning	-	+	++	++
Inter-period	–	-	-	+

Table 3.7.: Application degree of all three validation types.

consequently can answer **RQ 1.2** in affirmation. The same is true for the two hypotheses **H 1.3.1** and **H 1.3.2** regarding **RQ 1.3** that indicate an increase in the number of validations through an increased documented communication and knowledge about the new process structure (including expert knowledge). Overall, the significant increase of the data *timeliness* along with the improvement in the dimensions *completeness* and *consistency* have led to a strong increase in data quality. The standardization in data quality and organizational structure through the implementation forms the groundwork for the extensive data analyses in Chapter 3 to 5. In particular, the plan-actual comparison presented in Chapter 5 will enable us to present results for *data accuracy*.

Part III.

Business Characteristics Extraction

Chapter 4.

Planning Data Analyses

Financial planning processes as described in Section 3.1.1 are very time-consuming tasks. However, timeliness of final data is crucial for efficient risk management. In parallel, completeness, consistency and accuracy have to be assured throughout the complete process. The process redesign presented in the previous section established the prerequisites regarding performance and organizational structure necessary to focus on these three data related quality dimension. Based on reduced process runtime and the realization of a rule-based validation service it is now reasonable to put special attention on the detection of patterns in historical and actual data. The goal of this enhancement is the extension of rule-based validation for complex decision support. Therefore, this section provides detailed insights into the structure of planning data and establishes the basis for in depth investigations in the following two chapters. Section 4.1 introduces the underlying data and derives hypotheses necessary to answer **RQ 2.1** to **RQ 2.3** (Section 1.2.2). In Sections 4.2 and 4.3 the outcome of detailed numerical analyses on the characteristics of Benford's Law in financial planning data is presented and first conclusions are drawn whether Benford's Law is consistent with – or can easily be combined with – the concept of weak planning data efficiency.

4.1. Methodology

In analogy with the Sections 3.5.1 to 3.5.3 this methodology part starts with the underlying financial planning data sample and the preparation of all sub-samples. Based upon that, Section 4.1.2 describes the conducted interferential analyses in combination with the precision measure we introduce: the average fulfilment rate. These preparatory explanations and the content presented in Section 2.3.1 allow us to formulate the hypotheses to be evaluated throughout this chapter in Section 4.1.3.

4.1.1. Empirical Data Description

The data set to be evaluated is the cash flow-oriented financial planning data we have access to at our *industry partner*. In more detail, we have 25 data pools from May 2005 (2 2005) to June 2011 (2 2011) available in the first study (Section 4.1.2). In a second study (Section 4.1.3) progressing time and additional data deliveries enlarged our empirical data set to 27 data pools with two additional samples for September and November 2011 (3/4 2011). Each of these data pools reflects one data delivery and contains the planned values of all entities that have handed in planning data during this delivery. The planning data per entity reports expected issued and received invoices, cash flows, tax payments and so on for 15 months into the future. Initially, the financial planning data was delivered quarterly in the months of February, May, August and November. Due to internal re-structuring, the delivery months have changed to March, June, September and November since 2008. For the Benford evaluation, we take the complete set of planned values per delivery and subsidiary (entity) into account. Planned values are delivered quarterly by more than 100 subsidiaries. For the evaluation of weak planning efficiency, we compute the revisions per plan item and sort them according to their occurrence (first revision, second revision, etc.). Due to the data delivery on a quarterly basis with a maximum planning horizon of 15 months, each item is planned five times and we have four revisions. The efficiency calculation based on these four revisions for all data samples is finally based on the linear dependency between the revisions r_{t-1} and r_t.

To increase the robustness of our first results in 4.1.2, we conduct the evaluations for nine data samples: these are the complete data sample always including the entire data pool and 8 sub-samples including only parts of the data pool. The variety of samples provide different perspectives on the available data. More detailed, we first divide the basic set ([1] *complete*) into data delivered by [2] *large* and [3] *small entities*. Thereby, the size of the entity is determined by the number of planned values it delivers. This distinction results from the expectation that large entities with a high number of planned values can put more effort in their planning data generation due to economies of scale (Williamson, 1991) and, hence, achieve a higher quality. Since planning data has not been examined in literature before, this distinction is based on expert knowledge gained from interviews within the enterprise. The second perspective distinguishes data with [4] *positive* and [5] *negative* prefixes. This separation is based on observations made, for instance, by Carslaw (1988) and Thomas (1989), who observed different digit distributions for reported positive and negative numbers. The remaining four sub-samples are combinations from the distinctions listed above, that is [6] *positive large*, [7] *positive small*, [8] *negative large*, and [9] *negative small*.

For the advanced results in Section 4.1.3 we again divide the full data set *complete* into five disjoint sub-samples, determined by the division a specific plan item delivering entity belongs to. The five sub-samples (or groups) are *material science (MS)*, *crop science (CS)*, *health care (HC)*, *holding (HO)*, and *diverse (DV)*. The last category contains all entities that can not exactly be assigned to one of the divisions. This differentiation allows us to examine the influence of different parameters on our metrics in Chapter 6. A few examples are now described. On the one hand, *MS* supplies a huge amount of accessories to the automotive industry. Since the automotive industry strongly depends on macro economy, the macro-economic development is likely to impact financial planning in *MS*, too. This impact is caused by the close relationship between MS and the automotive industry producing vendor parts for this industry. Another example of external data that might impact planning processes of *CS* are particular weather conditions. On the other hand, according to experts within the company, *HC* has no explicit underlying influences but simply depends on the maturity of licenses. A long-term goal of our work is to relate group-differences with respect to quality metrics to external factors, if indicated.

Generally, the number of delivered planned values increased from $27,511$ in May 2005 to $72,141$ in June 2011. Additionally, the creation of the sub-samples as described above brings about highly different sample sizes. To tackle this issue and to create a solid and comparable basis for the interferential analyses (cp. Section 4.1.2), we conduct a *normalization* as follows: We set the minimum sample size N equal to the number of items included in the minimum of all delivery-specific sample sizes. Afterwards, we reduce all other sample sizes to N. To do so, we uniformly draw N items from the respective sample. If the reduced set is not significantly smaller than the original one, this reduction does not change the digit distribution, yet to avoid biases in case of a stronger reduction, the reduction is carried out multiple times and the average of the resulting distributions is calculated. To obtain a reasonable trade-off between evaluation performance and accuracy of the result, we performed a simulation study and determined the required number of reductions that have to be carried out considering the distribution of the original sample size and the distribution of the reduced sample size. For instance, we found out that in case of a reduction higher than 90% (e.g., the *complete* data sample size in June 2011 was $72,141$ which had to be reduced to the minimum number $5,225$), we have to calculate the average of 50 reductions.

To summarize, in order to secure the comparability of our results, the normalization of the sample size facilitates the creation of a solid data pool and, thus, robust results. However, due to structural limitations it was not possible to conduct a normalization between the subgroup data samples (very small sample size *HO*). Hence, the inter-subgroup results may in some way be slightly biased by the sample size. Nevertheless, the sample sizes were the same for weak planning efficiency as well as Benford's Law and therefore the indicated differences should be the same. The above-described data preparation (sub-samples and normalization) also includes the

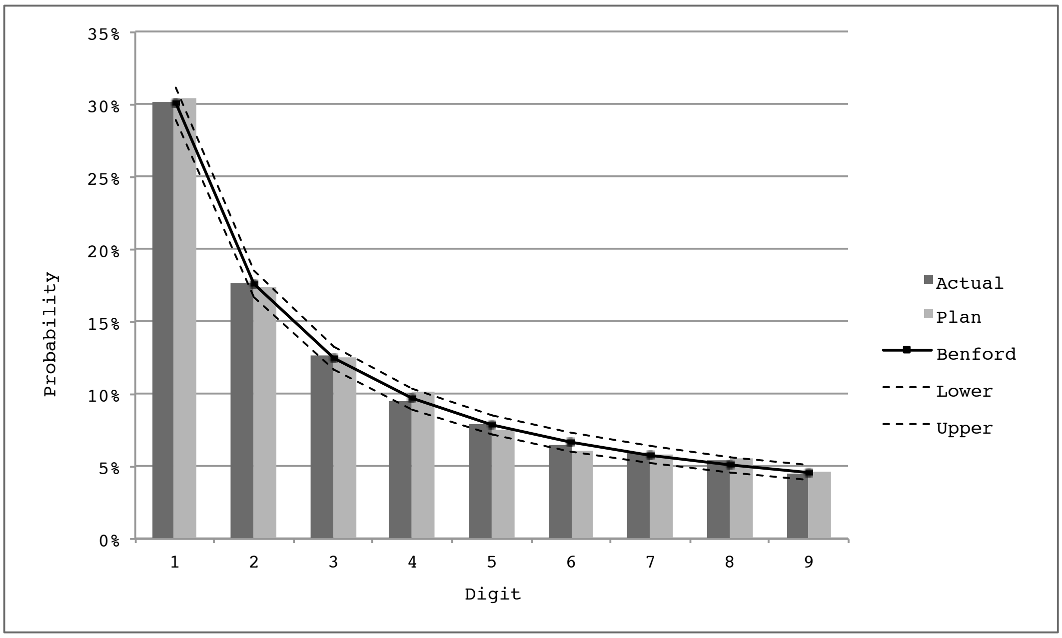

Figure 4.1.: First digit distribution: in actual data, planning data, and according to Benford's Law, including 95% confidence interval.

deletion of all planned values with an absolute nominal less than 10 (depends on the currency and includes zero values) to avoid procedural problems with Benford's Law according to Nigrini and Mittermaier (1997).

Before starting with the core evaluation of this chapter it is necessary to prove the conformity of actual data (for detailed description of available actual data please refer to Section 5.1.1) with Benford's Law. In doing so, we calculate the digit distributions in 2011 for *complete* actual data and find a nearly perfect match as can be seen in Figure 4.1 and 4.2. Especially for the first digit (4.1) we observe a conformance of 100% along with the non-significant chi-square value 1.43. Nearly the same is true for the second digit (4.2): conformance of 90% and chi-square value 2.85, respectively. In both figures we added the corresponding plan values, however, while the first digit results are already conform with actual data ($89\%/\ 9.02$), the second digit results show significant deviations ($60\%/\ 24.46$). The following sections include a detailed discussion of these effects.

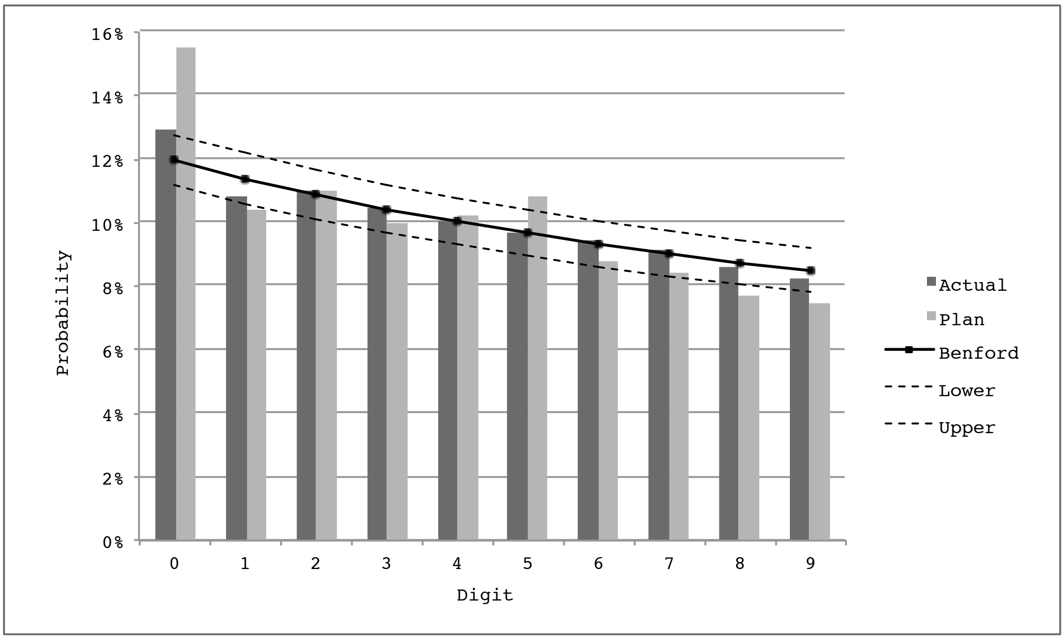

Figure 4.2.: Second digit distribution: in actual data, planning data, and according to Benford's Law, including 95% confidence interval.

4.1.2. Research Design

To date, related literature (cp. Section 2.3.1) consults two kinds of interferential analyses to investigate if data is conform to Benford's Law: interferential analyses for *(i)* single digits (e.g. *z-statistic*), and *(ii)* the complete distribution (e.g. *chi-square statistic*, *mean absolute deviation*). Yet, all of these approaches aim to detect significant deviations from the expected distribution yielded by Benford's Law. None of the approaches offers the possibility to make statements on the degree a distribution satisfies Benford's Law. The mean absolute deviation (*MAD*) only allows for a descriptive indication; the chi-squared test is suitable for a pure true-false view of the distribution without any differentiation in between. Moreover, the results of these analyses strongly depend on the size of the data sample. For instance, the width of a confidence interval is solely calculated based on the p-value and the number of observations.

The first goal of this section is the derivation of an alternative heuristic to overcome these shortcomings. Nevertheless, we apply *MAD* in later parts of this work since restrictions in the empirical data limit the applicability of the newly derived approach (cp. Section 5.1.2). Nigrini (2000) states that a desirable approach to investigate conformity to Benford's Law should fulfil the following requirements:

1. The test shall measure the conformity to the expected distribution, not only with single digits or digit combinations,
2. the result shall be independent of the sample size,
3. the test shall be implementable and understandable for users in practice, and
4. the conclusion of the test shall be objectively determinable.

Existing approaches fail to fulfil the independence requirement *2.* To fulfil *2.*, the normalization approach to eliminate the dependence on the basic sample size was presented in the previous section.

Independent of the application field, the main challenge for all inferential analyses is the interpretation of deviations from the expected Benford distribution. Nigrini and Mittermaier (1997) and Durtschi et al. (2004) apply single digit analyses based upon z-statistics and distribution analyses (chi-squared test). However, other scholars, e.g. Busta and Weinberg (1998), use neural networks to evaluate deviations from the expected distribution. The latter approach delivers better results than the above-mentioned digit analyses, however, to the disadvantage of type I errors which are an indication for fraud although data is correct (so-called *over auditing*). Since we want to keep the rate of type I errors low and to incorporate a more differentiated view of the fulfilment degree (cp. requirement *1.*), we introduce the average fulfilment rate (*AFR*) as a heuristic to measure the degree of confirmation between two distributions. The *AFR* of a data sample reflects the percentage of digits not deviating significantly from the expectation with respect to the digit position i. Accordingly, the calculation of the average fulfilment rate per distribution is based on the z-statistic per digit. For each digit, we calculate the 95% confidence interval in dependence of the basic p-value $.05$. Based upon this interval, we can decide if the digit probability significantly deviates from the expectation. In case of a deviation, we assign the digit with 0 and in case of conformity to 1. The *AFR* is then the mean of all decisions, for instance, in case of one digit deviating significantly in i, the AFR_i would be 88.9%. We can calculate AFR_i of the digit position i in the following way:

$$AFR_1\left(P_1\left(d\right)\right) = \frac{1}{9}\sum_{j=1}^{9} 1_{[l_j,u_j]}\left(P_1\left(j\right)\right), \tag{4.1}$$

$$AFR_2\left(P_2\left(d\right)\right) = \frac{1}{10}\sum_{j=0}^{9} 1_{[l_j,u_j]}\left(P_2\left(j\right)\right), \tag{4.2}$$

where $P_i\left(j\right)$ denotes the probability for each digit j in position i for the evaluated empirical data sample. $P_i\left(d\right) \in \left[0,1\right]^n$ denotes the vector of all digit probabilities with $P_1\left(d\right) = \left(P_1\left(1\right), \ldots, P_1\left(9\right)\right)$ and $P_2\left(d\right) = \left(P_2\left(0\right), \ldots, P_2\left(9\right)\right)$. $1_{[l_j,u_j]}$ is the indicator function which is 1 in the denoted interval $[l_j, u_j]$ and 0 else. For the confidence interval, the lower bound l_j and upper bound u_j are calculated separately for each digit j. To demonstrate the indication of the *AFR*, we compared the chi-square value and the *AFR* for the complete data sample. This examination reveals a highly significant dependence with $\tau = -.76$, $p < .001$. Indeed, particularly a chi-square value less than 17.53 (i.e. there is no significant deviation between the distributions based upon $p = .05$) is significantly correlated ($\tau = .52$, $p < .01$) to an AFR_1 of 77.8% (i.e. seven of nine digits do not deviate significantly). Hence, the *AFR* yields the same significance as the chi-square test, yet, firstly, it is much easier to understand and to interpret than the chi-square value. Secondly, it offers a differentiated indication of the analysed data sample's conformity to the expected distribution. According to Nigrini (2000), both properties are decisive advantages for the implementation into a decision support service (cp. *3.* and *4.* in the requirements list).

As mentioned earlier, in forecast literature (cp. Section 2.3.1) it is very common to test for weak-form forecast efficiency (cp. Batchelor (2007), Fildes and Stekler (2002), Lux (2009), Schuh (2001)). We adapt this method for data quality assessment. Following the original idea, we propose to test whether financial planning data follows no predicable patterns. Hence, we test if data revisions follow a random walk. Given that a future event (e.g. cash flow) is planned several times at different points in time, we can define a data revision $DR_{i,t}$ at point in time t for item i as

$$DR_{i,t} = PLAN_{i,t} - PLAN_{i,t-1}, \tag{4.3}$$

where $PLAN_{i,t}$ is the planned value for item i in t, and $PLAN_{i,t-1}$ the former value in $t-1$. Furthermore, to test for correlation we use the following OLS regression:

$$DR_{i,t} = B * DR_{i,t-1} + \epsilon. \tag{4.4}$$

We call this adaptation *weak-form planning efficiency*. For the evaluation we can aggregate the estimates on various levels and test whether groups follow certain inefficient planning patterns. We then aggregate the absolute estimate and denote it $|B|$.

For the statistical evaluation of the results generated based upon the *AFR* and $|B|$, we conduct two kinds of non-parametric interferential analyses: *(i)* Kendall's correlation, and *(ii)* the Wilcoxon signed-rank test. We decided to utilize non-parametric approaches since we found, as a result of a Shapiro-Wilk distribution test, a significant deviation from the normal distribution for most of the tested treatments (for a more detailed explanation of the treatments, please refer to Section 4.1.1). We opt for the Kendall correlation coefficient τ since we have a relatively small sample size ($N = 25/N = 27$) along with many tied ranks Field (2009). The Wilcoxon signed-rank test is chosen as it is the most common non-parametric test. According to Field (2009), we always add the 1-tailed level of significance p and the test statistic T (denoting the smaller value of the two rank sums) to the reported absolute value. Since we usually evaluate directed hypotheses it is reasonable to report the 1-tailed level of significance. Only exception of this notation are the numbers in Table 4.5. Here we do not evaluate directed hypotheses and consequently denote 2-tailed p-values. In order to back up the robustness of our analyses, we also calculate Pearson correlations and apply t-tests with similar results.

4.1.3. Hypotheses

As we have argued in Section 2.3.1, financial planning data meets the necessary criteria to follow Benford's Law due to its structure. Moreover, since actual data is conform with Benford's Law, it is reasonable to assume that planning data is also conform with Benford Law. Additionally, a major foundation of our evaluation is expert knowledge gained from our industry partner on the quality assurance measures carried out during the six year time period which spans our data sample. We can demonstrate that the output data of the financial planning process was continuously increased, due to process optimization and compliance enhancement. We can also state that none of the improvements were directly and explicitly related to pushing the data towards the Benford distribution. Based upon this determining factor, the conformity has to be verified between the digit distribution in the underlying financial planning data sample and the logarithmic distribution present in Benford's Law in order to answer research question **RQ 2.1** (*Does financial planning data follow Benford's Law?*). To do so, we set up the following hypothesis:

H 2.1.1 – EX-ANTE QUALITY METRIC –
AFR increases over the considered time period.

To extract robust evidence for **H 2.1.1**, formally, we investigate 18 sub-hypotheses: For each of the investigated nine data samples [1]-[9] (cp. Section 4.1.1), the conformity of the first and second digit of the underlying numbers to Benford's Law is tested. This consideration leads to 18 treatments to be tested. In more detail, these are: [1.1] complete first digit, [1.2] complete second digit, [2.1] large first digit, [2.2] large second digit, and so forth, ending with [9.1] negative small first digit and [9.2] negative small second digit.

In order to find evidence for **RQ 2.2** and hence the appropriateness of digital analyses in financial planning, we have to delve deeper into details of the data structure. In this work, we concentrate on *(i)* differences between reported negative numbers and positive numbers (Carslaw, 1988; Thomas, 1989) and *(ii)* large and small entities (as large companies, due to their size and their planning volume, may be able to put a greater effort as well as more expertise into the financial planning). Thus, we investigate the following hypotheses **H 2.2.1** and **H 2.2.2**:

H 2.2.1 – EX-ANTE QUALITY METRIC –
AFR in data delivered by large entities is higher than in data delivered by small entities.

H 2.2.2 – EX-ANTE QUALITY METRIC –
AFR in data with a positive prefix is higher than in data with negative prefix.

To validate these hypotheses, we again conduct analyses for the first and second digits in multiple data sub-samples: For **H 2.2.1**, these data samples are [1.1] complete first digit, [1.2] complete second digit, [4.1] positive first digit, [4.2] positive second digit and [5.1] negative first digit, [5.2] negative second digit. For **H 2.2.2**, we consult [1.1] complete first digit, [1.2] complete second digit, [2.1] small first digit, [2.2] small second digit and [3.1] large entities first digit, [3.2] large entities first digit. **H 2.2.1** and **H 2.2.2** potentially provide us with insights that we can include into a decision support service for financial planning managers.

In a further study we enrich the above-mentioned investigation by five additional sets based on the structure of the industrial partner (cp. 4.1.1) and extend the findings for **H 2.1.1**. However, this second investigation is not limited to the conformity with Benford Law, but also evaluates whether the metric *weak planning efficiency* is appropriate for application in the domain of financial planning. Since we are mainly interested in interdependencies between the two metrics, we conducted only the AFR_1 evaluation. By tracking the developments of both metrics over multiple years with intensive data quality management activities, we further evaluate **H 2.1.1** and addressed **RQ 2.2** in depth with **H 2.2.3**. Since the data samples do not only differ in the business model of the included subsidiaries, but also in organizational matters, it is reasonable to expect such differences. For instance, mergers and acquisitions as well as spin-offs have a great influence on the data structure. To see whether the perceived data improvement (cp. 1.3) is reflected in weak planning efficiency, the additional question **RQ 2.3** is addressed by two further hypotheses:

H 2.2.3 – EX-ANTE QUALITY METRIC –
AFR varies between lines of business.

H 2.3.1 – EX-ANTE QUALITY METRIC –
$|B|$ decreases over the considered time period.

H 2.3.2 – EX-ANTE QUALITY METRIC –
$|B|$ varies between lines of business.

For various tasks of a financial controller such as forecasting or individual feedback in the field of complex financial planning data, it is essential to have a detailed understanding of distinctive data characteristics. Hence, the results in this chapter are essential for research design in the following chapter and, finally, for the creation of a decision support service (6). In addition, we are interested in a comparison of results obtained with both metrics, and how they interrelate, as both are expected to indicate planning data quality. Since an increasing data quality is associated with a decreasing $|B|$ and an increasing AFR it is reasonable to expect a negative correlation between the two.

4.2. Benford's Law

This section presents a detailed examination of the first hypotheses set up in Section 4.1.3. To address **H 2.1.1**, we investigate the dependency between progressing time and the data's conformity to Benford's Law for the first and second digits in 9 data samples (cp. Section 4.2.1). In the following section, we generate detailed knowledge about the data characteristics in order to address **H 2.2.1** to **H 2.2.3**.

4.2.1. Trend Analyses and Robustness

In order to address **H 2.1.1**, we perform the trend analyses for all nine data samples described in Section 4.1.1, each of them for the first and second digit. Altogether, 18 treatments can thus be tested. The results of this evaluation for AFR_1 and AFR_2 are listed in Table 4.1. We calculated the mean over 25 deliveries along with the correlation τ and the significance level p. Although we chose the rather conservative Kendall's τ (the average absolute value is only around $2/3$ of a Spearman correlation according to Field (2009)) we found a medium or strong correlation for 15 of 18 treatments. The only non-significant results were discovered for *positive* data, although we found an at least small positive correlation even for [4.1] ($\tau = .19$), [6.1] ($\tau = .09$) and [7.2] ($\tau = .19$).

For AFR_1, the effect size ranges from $\tau = .44$, $p < .01$ in [2.1] and $\tau = .45$, $p < .001$ in [9.1] to $\tau = .59$, $p < .001$ in [1.1] and even $\tau = .65$, $p < .001$ in [3.1]. For AFR_2, in analogy with the general AFR level, the trends are alleviated compared to AFR_1. Nevertheless, they are all at least medium strong ($\tau = .27$, $p < .05$ in [3.2] or $\tau = .28$, $p < .05$ [2.2]) and even range to strong dependencies in [5.2] ($\tau = .50$, $p < .001$) and [1.2] ($\tau = .51$, $p < .001$).

To ease reading, we show graphs of the development for two exemplary data samples: Figure 4.3 shows the AFR_1 (for all 25 deliveries); for [1.1], i.e. complete on the left hand side, and for [3.1], i.e. small on the right hand side. As easily can be seen, for [1.1], the AFR_1 clearly increases over time from a minimum value of 66.7% to 100% in the last four deliveries. A highly significant correlation between time and AFR_1 ($\tau = .59$, $p < .001$) is present. In [3.1], the trend itself is even stronger with $\tau = .65$, $p < .001$, yet, the mean AFR_1 over all 25 deliveries is higher in [1.1] (87.2%) than in [3.1] (70.7%). These differences provide first indications for the evaluation of **H 2.2.1** in the next section.

The difference described above can also be observed in the mean AFR_2 (29.6% in [1.2] and 10.4% in [3.2] data). Yet, as clearly demonstrated in Figure 4.4, the degree

Data sample	AFR_1			AFR_2		
	Mean	τ	p	Mean	τ	p
[1] Complete	87.2%	.59	.000***	29.6%	.51	.000***
[2] Large	81.4%	.44	.002**	57.6%	.36	.010**
[3] Small	70.7%	.65	.000***	10.4%	.27	.047*
[4] Positive	79.2%	.19	.102	48.4%	.44	.002**
[5] Negative	79.3%	.54	.000***	20.0%	.50	.001***
[6] Positive Large	70.0%	.09	.289	67.6%	.32	.020*
[7] Positive Small	60.6%	.53	.000***	26.8%	.19	.114
[8] Negative Large	64.9%	.46	.001***	47.2%	.28	.033*
[9] Negative Small	60.6%	.45	.001***	10.8%	.42	.004**

Table 4.1.: Evaluation results (Mean/correlation τ/p-value): AFR trend analyses for first and second digit position (AFR_1 and AFR_2) in nine data samples ($p^* < .05$; $p^{**} < .01$; and $p^{***} < .001$).

of conformity to Benford's Law in financial planning data is on a lower level in the second digit than in the first digit. Still, we observe a strongly positive trend in [1.2] ($\tau = .51$, $p < .001$) and at least a medium positive trend [3.1] ($\tau = .27$, $p < .05$).

Altogether, we are able to identify a significant positive trend in AFR_1 and AFR_2 for 15 of 18 data samples. For the data samples [4.1], [6.1] and [7.2] we found at least a small correlation (cp. Field (2009)). Based upon these findings, we can confirm **H 2.1.1**. Furthermore, the results clearly point to a difference in the quality indication of AFR_1 and AFR_2: increased data quality is likely to first lead to an increased AFR_1 and, in a second step, leads to an increased AFR_2. That is why we (i) observe a generally smaller AFR_2 throughout all samples, and (ii) AFR_2 is higher in data samples with a higher AFR_1. However, the robust positive trend is present both for AFR_1 and AFR_2, which verifies **H 2.1.1**. Although the average AFR_2 in our data set is not (yet) conform to the expected distribution, there is a strong indication for the conformity of high quality financial planning data to Benford's Law. Thus, **RQ 2.1** can be confirmed. Finally, these findings are extended by the results presented in Section 4.3, whereby this result separation is due to a different structure in the two studies – longer time horizon (second half of 2011) along with reduced evaluation (only AFR_1).

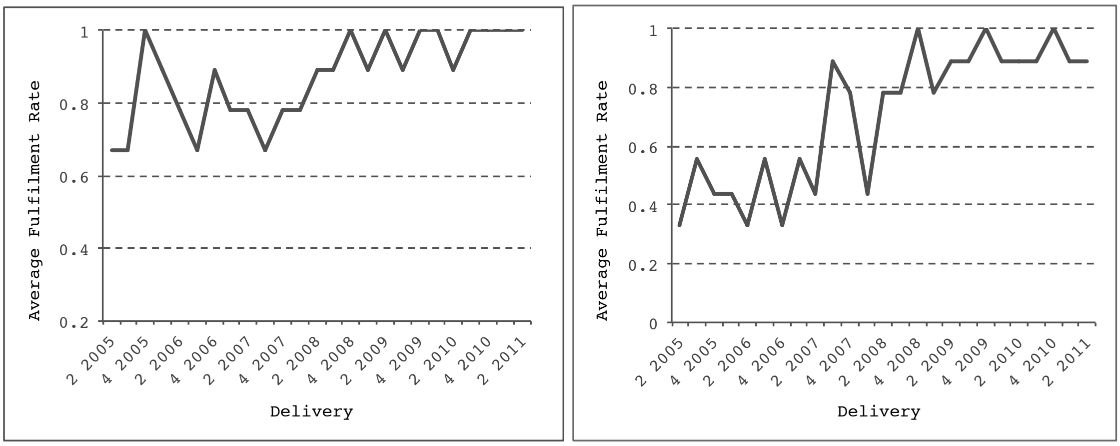

Figure 4.3.: AFR_1 development in the complete (on the left) and in the small data sample (on the right) over all deliveries (from 2 2005 to 2 2011).

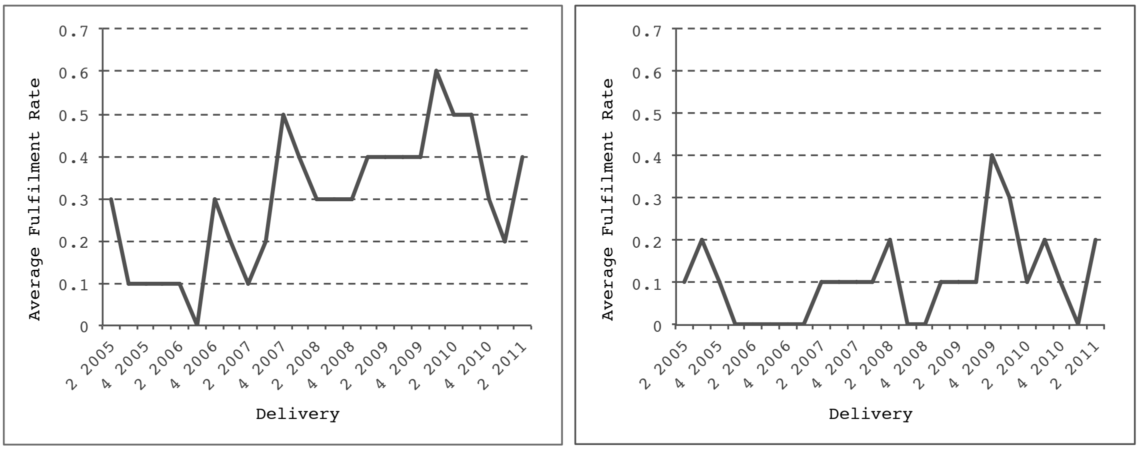

Figure 4.4.: Development of AFR_2 in the complete (on the left) and in the small data sample (on the right) over all deliveries (from 2 2005 to 2 2011).

4.2.2. Group Analyses

To validate **H 2.2.1** and **H 2.2.2**, we again conduct analyses for multiple data samples to assure either the robustness of the results if they are unique or to detect characteristics of the data clusters if the results are controversial. In case of **H 2.2.1** the investigated data samples are *complete*, *positive* and *negative*, both for the first and the second digit. Furthermore, we investigate the *complete*, *small* and *large* data sample to validate **H 2.2.2**, again for both digit positions.

The results for the investigation of **H 2.2.1** are listed in Table 4.2. We were able to show significant differences between small and large entities for 5 of 6 investigations. The differences in AFR_1 are significant for the *complete* data with 10.7 percentage points ($T = 2.32$, $p < .05$) and *positive* data with 9.4 percentage points ($T = 2.16$, $p < .05$), yet not for negative data (3.3 percentage points). The largest differences can be observed in AFR_2. Here, we observe significant differences in all samples: 47.2 percentage points in *complete* ($T = 12.12$, $p < .001$), 40.8 percentage points in *positive* ($T = 10.54$, $p < .001$), and 36.4 percentage points in *negative* ($T = 10.11$, $p < .001$).

Table 4.3 contains the evaluation results for AFR_1 and AFR_2 with respect to **H 2.2.2**. Interestingly, in none of the data samples *complete*, *large*, and *small*, a significant difference between positive and negative numbers can be shown for AFR_1. In contrast, in AFR_2 we find a highly significant deviation for all data samples. The difference varies from 16.0 percentage points to 28.4 percentage points, $T = 7.53$, $T = 6.51$ and $T = 5.06$, $p < .001$ in the complete, positive and negative data sample. Of note, however, these findings are completely conform with the results of, for instance, Carslaw (1988) and Thomas (1989), who described different rounding behaviour in dependency of the prefix, observable through patterns in the second digit distributions.

To summarize, we can confirm **H 2.2.1** for AFR_1 and AFR_2. The only non-significant difference (negative) has the correct direction, too. However, **H 2.2.2** can only be confirmed for AFR_2. Nevertheless, the results of this section are very important for the design of a business intelligence service based on digital analyses: For an application

Data sample	AFR_1			AFR_2		
	[Sample] Mean		p	[Sample] Mean		p
	Large	Small		Large	Small	
Complete	[2.1]81.4%	[3.1]70.7%	.015*	[4.2]57.6%	[3.2]10.4%	.000***
Positive	[6.1]70.0%	[7.1]60.6%	.021*	[6.2]67.6%	[7.2]26.8%	.000***
Negative	[8.1]64.9%	[9.1]60.6%	.207	[8.2]47.2%	[9.2]10.8%	.000***

Table 4.2.: Results (mean/p-value) of the AFR comparison between large and small data for first and second digit in the complete, positive, and negative data sample ($p^* < .05$; $p^{**} < .01$; and $p^{***} < .001$).

Data sample	AFR_1			AFR_2		
	[Sample] Mean		p	[Sample] Mean		p
	Positive	Negative		Positive	Negative	
Complete	[4.1]79.2%	[5.1]79.3%	.490	[4.2]48.4%	[5.2]20.0%	.000***
Large	[6.1]70.0%	[8.1]64.9%	.196	[6.2]67.6%	[8.2]47.2%	.000***
Small	[7.1]60.6%	[9.1]60.6%	.500	[7.2]26.8%	[9.2]10.8%	.000***

Table 4.3.: Results (mean/p-value) of the AFR comparison between positive and negative data for first and second digit in the complete, large, and small data sample ($p^* < .05$; $p^{**} < .01$; and $p^{***} < .001$).

of the AFR_1 as an error indicator only the company size is crucial and the sign can be ignored, whereas AFR_2 is sensitive to company size and sign. Furthermore, according to the findings of Section 4.2.1, AFR_2 should only be applied to data samples of large entities.

4.3. Planning Data Efficiency

Based on the findings of the previous section, we now present analyses of the quality metric *weak planning efficiency* (cp. Section 2.3.1) along with an extended evaluation of the conformity with Benford's Law. Furthermore, this section relates the findings of both newly introduced quality metrics and investigates differences and similarities.

4.3.1. Trend Analyses and Robustness

For the analyses conducted in this chapter we rely on the data basis of the five subgroups as aforementioned. To ease reading, we refer to the first digit conformity AFR_1 as AFR in the this section, since we do not evaluate the second digit distributions here. To increase the robustness of the findings for **H 2.1.1** and to address **H 2.3.1** we analyse the trend in AFR and $|B|$ over time for each data sample. The results listed in Table 4.4 are clustered in: *(i)* conformity with Benford's Law (AFR), *(ii)* weak planning efficiency ($|B|$) and *(iii)* linear correlation between *(i)* and *(ii)*. For *(i)* and *(ii)*, we report the mean over the 25 deliveries along with the non-parametric correlation τ and the significance level p. Although we choose the conservative Kendall's τ, we find a medium or strong AFR-correlation for five out of our six groups. The only non-significant trend is discovered for DV. Even for DV we find a small positive correlation ($\tau = .19$). For weak efficiency we obtain more heterogeneous results. We observe the expected significant decrease ($\tau = -.25$, $p < .05$) for CS. However, we also observe the opposite relationship with a significant and strong positive trend in $|B|$ for *HC* ($\tau = .39$, $p < .01$).

As an example, the graph on the left-hand side of Figure 4.5 shows the AFR for all 25 deliveries; the AFR increases over time from 66.7% to 100% throughout the final four deliveries. We find similar behaviour in all six subgroups, which again supports **H 2.1.1**. Although the decreasing $|B|$ depicted in the right pictorial in 4.5 indicates a positive trend in planning efficiency and supports **H 2.3.1**, unfortunately the results in the other groups do not support **H 2.3.1**. Hence, we cannot accept **H 2.3.1**.

A first explanation for this divergence might be expressed by the high standard deviation of $|B|$ (up to 19% for *MS*). Although the standard deviation varies among subgroups, it is generally high compared to the mean values. This uncertainty might be a result of the multiple impacts driving the planning data generation and requires a further differentiation with regard to the distinctive characteristics of subsidiaries as aforementioned.

Data sample	Conformity AFR			Weak Efficiency $\lvert B \rvert$			Dependency	
	Mean	τ	p	Mean	τ	p	τ	p
MS	44.8%	.57	.000***	0.385	.10	.235	.08	.309
CS	83.7%	.61	.000***	0.401	$-.25$	.042*	$-.21$	.084
HC	82.8%	.43	.003**	0.575	.39	.003**	.13	.207
HO	67.8%	.30	.028*	0.586	.16	.136	.28	.042*
DV	52.0%	.19	.102	0.276	.04	.390	.05	.378

Table 4.4.: Evaluation results (mean/correlation τ/p-value): AFR and $|B|$ trend and dependency analyses in five data samples ($p^{*} < .05$; $p^{**} < .01$; and $p^{***} < .001$).

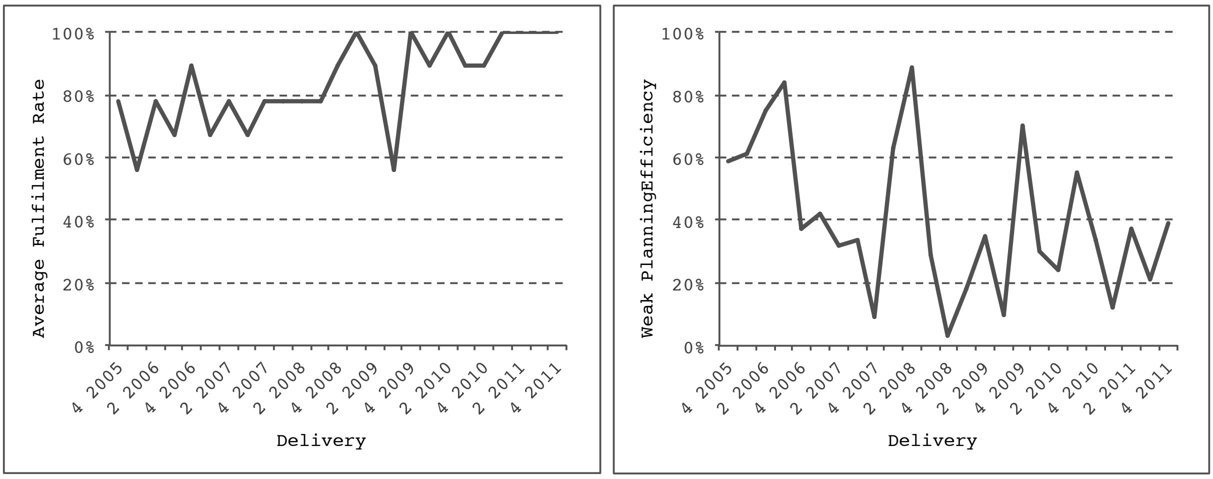

Figure 4.5.: Development of AFR (on the left) and of $|B|$ (on the right) in the *CS* data sample over all deliveries (from 4 2005 to 4 2011).

Consequently, as shown in the right two columns of Table 4.4, the two metrics seem to represent different aspects of the data. On the one hand, we find a negative dependency ($\tau = -.21$, $p = .084$) in *CS* data. On the other hand, we observe a positive one in *HC* data ($\tau = .13$, $p = .207$) and even a significant positive one in *HO* data ($\tau = .13$, $p < .05$).

4.3.2. Group Analyses

To validate the final two hypotheses of this chapter, **H 2.2.3** and **H 2.3.2**, Table 4.5 shows our results for the above-described non-parametric tests on group differences

Data sample	Cross-sample test p-values [Conformity with Bendford's Law (AFR)/ weak planning efficiency ($\|B\|$)]				
	MS	CS	HC	HO	DV
MS	$1.00/$ 1.00	$.000^{***}/$ $.778$	$.000^{***}/$ $.201$	$.000^{***}/$ $.003^{**}$	$.106/$ $.098$
CS	$.000^{***}/$ $.778$	$1.00/$ 1.00	$.604/$ $.493$	$.001^{***}/$ $.042^{*}$	$.000^{***}/$ $.036^{*}$
HC	$.000^{***}/$ $.201$	$.604/$ $.493$	$1.00/$ 1.00	$.000^{***}/$ $.288$	$.000^{***}/$ $.065$
HO	$.000^{***}/$ $.003^{**}$	$.001^{***}/$ $.042^{*}$	$.000^{***}/$ $.288$	$1.00/$ 1.00	$.002^{**}/$ $.004^{**}$
DV	$.106/$ $.098$	$.000^{***}/$ $.036^{*}$	$.000^{***}/$ $.065$	$.002^{**}/$ $.004^{**}$	$1.00/$ 1.00

Table 4.5.: Group differences with respect to AFR and $|B|$ ($p^{*} < .05$; $p^{**} < .01$; and $p^{***} < .001$).

for both metrics. The results are denoted in a cross-group structure (25 combinations) whereby the depicted p-values within a subgroup are trivial ($1.00/1.00$ for $AFR/|B|$).

We find highly significant differences for almost all AFR sample tests with p-values close to zero except for the samples *MS* and DV. Again, we observe more heterogeneous results for the $|B|$ sample tests. In contrast to our trend evaluation, we do not find a strongly deviating behaviour of *CS* but rather two groups of test results: the pairwise differences of *MS, CS, HC* are insignificant, while differences between *HO* or *DV* and each of the samples in *MS, CS, HC* clearly reveal differences.

In summary, we can confirm **H 2.1.1** and **H 2.2.3**, while we cannot support **H 2.3.1** and **H 2.3.2**. The heterogeneous results in the weak planning efficiency evaluation for the hypotheses **H 2.3.1** and **H 2.3.2** require further discussion and more in-depth investigation. They indicate more complex relationships between Benford's Law and weak efficiency in financial planning data. Interestingly, the data groups in *MS, CS, HC* contain only "real" subsidiaries within the same lines of business, whereas the other two are more heterogeneous. For example, entities in *DV* cannot clearly be assigned to one line of business. Although these phenomena cannot be explained until now, at least these distinctive characteristics are somewhat reflected by the results of our cross-group comparisons.

Chapter 5.

Data Quality Metrics

The previous chapter introduced two different quality measures applied to financial planning data. Since the weak planning efficiency results revealed no systematic behaviour, the present evaluation focusses on the quality indication provided by a conformance with the expected Benford distribution. Nevertheless, weak planning efficiency can be applied for case-sensitive investigations as it is shown in Section 6.3.1. To assure the validity of a Benford-based ex-ante quality indicator, a benchmark is required. Consequently, this chapter introduces the most frequently used quality indicator data accuracy (cp. Section 2.3.2) and evaluates its ability to reflect perceived data quality. Furthermore, we investigate the relationship between accuracy and Benford's Law to prove the similar quality indication. The foundation for the benchmark evaluation is provided in Section 5.1 with a detailed description of the complex data base underlying the calculation and its preparation. Based upon this methodology, Section 5.2 presents the core evaluation results and includes different dimensions of data accuracy, their indications, and relationships to the previously introduced indicators. Finally, this section concludes with first results for **RQ 3.1** representing the foundation for the steps towards a decision support service carried out in the following chapter.

5.1. Methodology

In analogy to previous chapters, this *methodology* section comprises of a detailed data and analyses description along with the hypotheses that should be evaluated in this chapter. Although, the available data in this chapter is mostly the same as before, important changes had to be made to enable the plan-actual comparison (Section 5.1.1). These adaptations also led to slightly different evaluation methods as described in Section 5.1.2.

5.1.1. Empirical Data Description

In contrast to the previous chapter, the calculations in this chapter are based on parts of the planning data. According to Section 4.1.2, each data sample for the digit distribution calculation contained all items of one delivery, comprising of 15 months. However, since each of these months have a different value-date and the value-date is essential for accuracy (plan-actual) calculation, it was necessary to split up the data. Figure 5.1 illustrates this dependency for value and plan quarters. In that way, for each quarter of actual data there exist five quarters of planning (due to the 15 months planning horizon and the resulting overlap between the deliveries– cp. also Section 4.1). Accuracy values can now be calculated for combinations of actual values with corresponding planned values per plan quarter. Each of these five accuracy values is characterized by the planning horizon of the underlying planning data lasting from $1 - 3$ to $13 - 15$ months. Furthermore, the actual data bases is limited to invoices issued and received for the time period March 2008 (1 2008) to March 2012 (1 2012). These restrictions result from technical limitations within our industrial partner for the availability of actual data. To assure compliance, only relative values and linear transformations of actual values are included in this work.

Similarly to all other evaluations, we again conducted the analyses for numerous subsamples. The definition is in conformance with previous chapters, however, we limited the evaluation to three main samples *Complete*, *Large*, and *Small* along with four

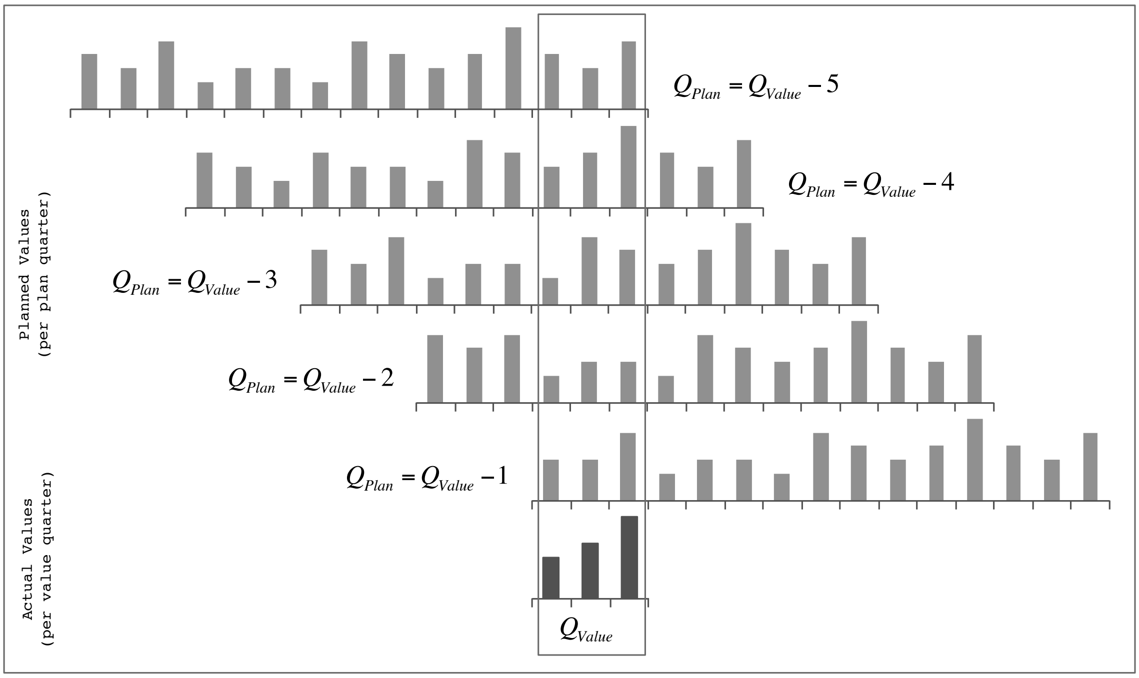

Figure 5.1.: Planned and actual values underlying accuracy calculation in rolling data.

of five sub-group samples (*HC*, *MS*, *CS*, and *DV*). This restriction is valid since the expectation for differences between positive and negative data resulted from digital analyses (Carslaw, 1988) that are not in the focus of this chapter. Furthermore, holding planning data entirely reflects subsidiary planning data and hence planning errors are determined by errors in subsidiary data. Therefore, we exclude holding data from accuracy evaluation. Nevertheless, we should be able to generate robust results in this chapter based on the investigation of the remaining seven sub-samples.

5.1.2. Research Design

One major effect of the reduced data basis described in the previous section is the limitation of the Benford analyses to the mean absolute deviation (MAD) calculation. The reductions lead to a relatively low number of plan items for the different samples causing wide confidence intervals in interferential analyses. That again would lead to perfect AFR-results for nearly all data samples throughout the complete time period and hence provides no further insights. With the application of the metric MAD we loose the significance indication, however, we gain a more detailed differentiation over time that is more important in the context of the analyses performed in this chapter. Furthermore, we are only interested in the relationship between first digit distributions and accuracy and consequently calculate MAD as follows:

$$MAD = \frac{1}{9} \sum_{i=1}^{9} |p^{exp}(i) - p^{obs}(i)|, \tag{5.1}$$

where $p^{exp}(i)$ denotes the expected and $p^{obs}(i)$ the observed probability for digit i.

The accuracy assessment of the planning data first requires the calculation of the planned foreign exchange exposure E^P and the corresponding actual exposure E^A. E^A is calculated based on booked transactions and consequently represents no real exposure. However, to ease reading, both values are denoted as exposure. To avoid mapping effects, we calculate the absolute percentage planning error up from the lowest accessible level- per currency within each entity. The following aggregation leads to three steps for the calculation per sub-sample:

$$APE_{Cur} = \frac{AE_{Cur}}{|E^A_{Cur}|} = \frac{|E^P_{Cur} - E^A_{Cur}|}{|E^A_{Cur}|}, \tag{5.2}$$

$$APE_{Ent} = \frac{AE_{Ent}}{E^A_{Ent}} = \sum_{Cur} \frac{|E^A_{Cur}|}{E^A_{Ent}} APE_{Cur}, \tag{5.3}$$

$$APE_{Group} = \frac{AE_{Group}}{E^A_{Group}} = \sum_{Ent} \frac{|E^A_{Ent}|}{E^A_{Group}} APE_{Ent}, \tag{5.4}$$

where the absolute error AE per currency in Equation 5.2 is calculated as the absolute difference between planned exposure E^P and actual exposure E^A. Further aggregated exposures and absolute errors AE are then calculated as simple sums: $E^A_{Ent} = \sum_{Cur} E^A_{Cur}$, $E^A_{Group} = \sum_{Ent} E^A_{Ent}$, $AE_{Ent} = \sum_{Cur} AE_{Cur}$, and $AE_{Group} = \sum_{Ent} AE_{Ent}$. Based upon that, the absolute percentage error APE is calculated as the ration of AE and corresponding actual exposure E^A– per currency in case of Equation 5.2. For the interpretation of the achieved results, it is important to note that the accuracy aggregation on entity and group level is a weighted aggregation. As can be seen in the equation chains 5.3 and 5.4, the calculation of the APE through absolute error and exposure is equivalent to a weighted mean of the $APEs$ one aggregation level below. These weights lead to an increased influence of planning errors with a high nominal E^A. Such a weighted aggregation is crucial to achieve results with practical impact. Moreover, this accuracy measure is a rather pessimistic indicator in the context of our industrial partner. Since accuracy is calculated separately for each entity (Equation 5.3) inter-company errors in planning data are taken into account twice and netting effects resulting from positive and negative errors are avoided completely. In addition, the planning currency differs from the booking currency in some special cases, for instance, due to legal requirements. We are aware of this special data structure and take it into account for our evaluation. However, we might not know all cases and that again can lead to increased planning errors.

Despite all these effects, the achieved results reveal a high overall accuracy in the financial planning data of our industrial partner as can be seen in the following chapters. Finally, the calculation of the percentage error PE is performed analogously:

$$PE_{Cur} = \frac{ER_{Cur}}{|E^A_{Cur}|} = \frac{E^P_{Cur} - E^A_{Cur}}{|E^A_{Cur}|}, \tag{5.5}$$

$$PE_{Ent} = \frac{ER_{Ent}}{E^A_{Ent}} = \sum_{Cur} \frac{|E^A_{Cur}|}{E^A_{Ent}} PE_{Cur}, \tag{5.6}$$

$$PE_{Group} = \frac{ER_{Group}}{|E^A_{Group}|} = \sum_{Ent} \frac{|E^A_{Ent}|}{E^A_{Group}} PE_{Ent}, \tag{5.7}$$

where the aggregated exposures and errors are again calculated as sums according to Equations 5.3 and 5.4. Per construction and due to netting effects, it holds $|PE| < |APE|$ for all sub-samples.

The decision for a different conformity indication in digital analyses and the additional accuracy calculation also influence the utilized statistical analyses: while we conducted the non-parametric Kendall's correlation before, we now chose the non-parametric Spearman correlation. Again, the decision for non-parametric approaches was based on results of a Shapiro-Wilk distribution test where we observed significant deviation from the normal distribution for most of the tested treatments. In addition, the new evaluations eliminated tied ranks and hence allowed us to apply Spearman rho (cp. Field (2009) and Section 4.1.2). For explanations of the reported 1-tailed level of significance p and the test statistic T (denoting the smaller value of the two rank sums) in addition to the reported absolute value please refer to Field (2009). Analogously to previous analyses, our results were backed up with Pearson correlations.

5.1.3. Hypotheses

The first goal of this chapter is to validate data accuracy as a data quality indicator in the domain of financial planning. The bases for this evaluation are unstructured expert interviews within our industrial partner revealing numerous quality assurance measures in financial planning: organisational restructuring, IS support for data delivery, process improvement through automation, integration of validation services, and workshops introducing the importance of appropriate planning to the subsidiaries among other measures. All this effort leads to a strongly improved perceived data quality over the last five years. Due to structural reasons, we are only able to evaluate the development of data accuracy since January 2008. Nevertheless, we address **RQ 2.4** (*Is data accuracy conform with perceived data quality?*) through the following two hypotheses:

H 2.4.1 – EX-ANTE QUALITY METRIC –
Data accuracy increases over the considered time period.

H 2.4.2 – EX-ANTE QUALITY METRIC –
Data accuracy increases with a reduced forecasting horizon.

While **H 2.4.1** results from the above-described increased perceived data quality, the second hypothesis is driven by the expectation of higher security with shorter planning horizon. In this vein, a reduced forecast horizon for the same forecast items should result in increased security and hence in increased data accuracy in good financial planning data.

In the next step, we aim to introduce Benford's Law as an ex-ante data quality indicator. To do so, two-fold conditions have to be fulfilled: *(i)* Benford's Law has to reflect perceived data quality similarly to data accuracy, and *(ii)* a dependency between high accuracy and small deviation from Benford's distribution has to be shown. Consequently, we are able to examine **RQ 2.5** (*Does Benford's Law in financial planning data assess data quality?*) based on three hypotheses:

H 2.5.1 – EX-ANTE QUALITY METRIC –
Conformity with Benford's Law increases over the considered time period.

H 2.5.2 – EX-ANTE QUALITY METRIC –
Conformity with Benford's Law increases with a reduced forecasting horizon.

H 2.5.3 – EX-ANTE QUALITY METRIC –
Conformity with Benford's Law is positively correlated with data accuracy over the considered time period.

To evaluate the hypotheses **H 2.4.1**, **H 2.5.1** and **H 2.5.3** we calculate the dependency between the indicators for all five planning horizons. This procedure results in a total sample size of $N = 85$ although we only have 17 points in time and hence strongly increases the robustness of the results. For the remaining hypotheses **H 2.4.2** and **H 2.5.2** we do the evaluation the other way round: five points in time for 17 sub-samples again leads to robust results with a total $N = 85$.

5.2. Benchmarking Benford's Law

In this section we answer the research question regarding the perceived data quality in financial planning: (i) we investigate whether data accuracy is conform with all expectations (5.2.1), and (ii) whether expectations are reflected in the conformity with Benford's Law (5.2.2). Finally, dependency between the two metrics is analysed.

5.2.1. Data Accuracy in Financial Planning

Data quality is a rather complex term. However, perceived data quality characteristics in financial planning can be brought back to the two effects expressed in **H 2.4.1** and **H 2.4.2**. In accordance to previous chapters, we again conduct our analyses to numerous sub-samples (cp. Section 5.1.2) to achieve robust results. The complete results for the development over time for accuracy (APE), the deviations from the expected Benford distribution (MAD), and the dependency are listed in Table 5.1.

The strongest accuracy improvement over time can be observed in *HC* with $\rho = -.73$, $p < .001$. Also the DV data ($\rho = -.33$, $p < .001$) and all aggregated samples ($\rho = -.46/-.39/-.41$, $p < .001$) reveal a strong negative correlation. However, the *MS* data exhibits no trend at all $\rho = .02$, $p = .419$. An explanation here might be the strong insecurity in macro-economy over last years (cp. Section 6.2.1). In *CS* the trend even is the other way round with a medium positive correlation $\rho = .13$, $p = .115$ probably caused by strong seasonal effects (cp. Section 6.2.2).

Table 5.2 contains the complete results for the dependency on the planning horizon, similarly structured as Table 5.1. However, the PE accuracy values are added to investigate systematic over-planning. In DV a reduced planning horizon has the greatest influence on data accuracy $\rho = .27$, $p < .01$. All other samples also exhibit medium to strong positive relationships, even if the *MS* effect is not significant $\rho = .13$, $p = .120$. It seams that the effect is stronger for large $\rho = .21$, $p < .05$ than for small $\rho = .16$, $p = .075$ subsidiaries. However, *CS* again differs strongly from the expectations $\rho = .04$, $p = .350$. For PE the trend results are quite homogeneous, however,

Data sample	APE			MAD Benford			Dependency	
	Mean	ρ	p	Mean	ρ	p	ρ	p
Complete	58.7%	$-.46$	$.000^{***}$	0.0049	$-.52$	$.000^{***}$	$.13$	$.126$
Large	54.7%	$-.39$	$.000^{***}$	0.0065	$-.48$	$.000^{***}$	$.16$	$.075$
Small	69.8%	$-.41$	$.000^{***}$	0.0072	$-.25$	$.010^{**}$	$.16$	$.071$
HC	44.4%	$-.73$	$.000^{***}$	0.0099	$-.42$	$.000^{***}$	$.47$	$.000^{***}$
MS	55.3%	$.02$	$.419$	0.0213	$-.66$	$.000^{***}$	$-.22$	$.023^{*}$
CS	84.2%	$.13$	$.115$	0.0122	$-.13$	$.114$	$-.08$	$.240$
DV	90.2%	$-.33$	$.001^{***}$	0.0076	$-.51$	$.000^{***}$	$.32$	$.002^{**}$

Table 5.1.: Evaluation results (mean/correlation ρ/p-value): Data accuracy APE and Benford MAD trend and dependency analyses over time in seven data samples ($p^{*} < .05$; $p^{**} < .01$; and $p^{***} < .001$).

Data sample	APE			MAD Benford			PE		
	Mean	ρ	p	Mean	ρ	p	Mean	ρ	p
Complete	58.7%	$.19$	$.043^{*}$	0.0049	$.30$	$.003^{**}$	13.0%	$.18$	$.055$
Large	54.7%	$.21$	$.027^{*}$	0.0065	$.40$	$.000^{***}$	20.6%	$.10$	$.176$
Small	69.8%	$.16$	$.075$	0.0072	$.38$	$.000^{***}$	-2.4%	$.12$	$.133$
HC	44.4%	$.19$	$.045^{*}$	0.0099	$.37$	$.001^{***}$	17.2%	$.12$	$.143$
MS	55.3%	$.13$	$.120$	0.0213	$.31$	$.002^{**}$	18.7%	$.10$	$.175$
CS	84.2%	$.04$	$.350$	0.0122	$.05$	$.326$	51.0%	$.05$	$.330$
DV	90.2%	$.27$	$.006^{**}$	0.0076	$.41$	$.000^{***}$	-19.5%	$.04$	$.355$

Table 5.2.: Evaluation results (mean/correlation ρ/p-value): Data accuracy APE/ PE and Benford MAD trend analyses over planning horizons in seven data samples ($p^{*} < .05$; $p^{**} < .01$; and $p^{***} < .001$).

the level of over- or under-planning reveals a strong heterogeneity. In this vein, we observe a maximum correlation in *Complete* of $\rho = .18$, $p = .055$ and nearly no development in *CS* and *DV* ($\rho = .05$, $p = .330$ and $\rho = .04$, $p = .355$). Nevertheless, none of the developments is significant. The most striking over-planning of 51% is in *CS* data and hence the PE results are completely in line with our previous findings. Moreover, they support our investigation of the over-planning behaviour in *CS* data in the following chapter.

As an example for the two trends discussed above, Figure 5.2 depicts the APE and MAD development for the fifth *Large* delivery. The graph shows the trend in APE and MAD for all 14 of 17 deliveries – this reduction is due to technical restrictions dur-

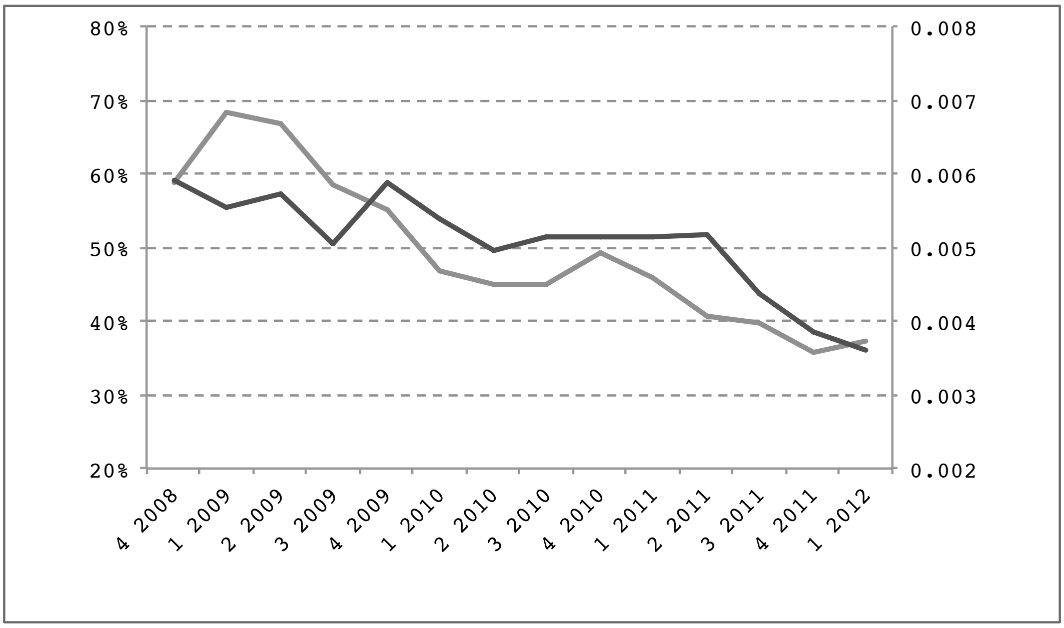

Figure 5.2.: Trend in accuracy APE (light grey) and Benford MAD (dark grey) for the *Large* data sample over all deliveries (1 2008 to 1 2012).

ing the trend extraction (for a detailed description of trend calculation refer to Section 6.1.2); APE decreases from an average of 65% in 2008 to less than 40% throughout the final deliveries. Figure 5.3 illustrates the influence of the planning horizon on APE and MAD in *Large* data. For data with a planning horizon of more than 12 months we observe an average planning error of more than 60%, meanwhile the planning error for data with planning horizon below 3 months is below 50%.

Summing up the different evaluations of this section, the evaluation of both hypotheses **H 2.4.1** and **H 2.4.2** revealed similar behaviour in five of our seven subgroups. Consequently, although the results support the hypotheses in most parts we can only conditionally answer **RQ 2.4** in affirmation. Moreover, the noticeable results for *MS* and *CS* data have to be further investigated regarding, for instance, external effects.

5.2.2. Data Accuracy Indication of Benford's Law

According to the previous section, data accuracy reflects the perceived data quality in financial planning data at large. However, data accuracy is an ex-post quality

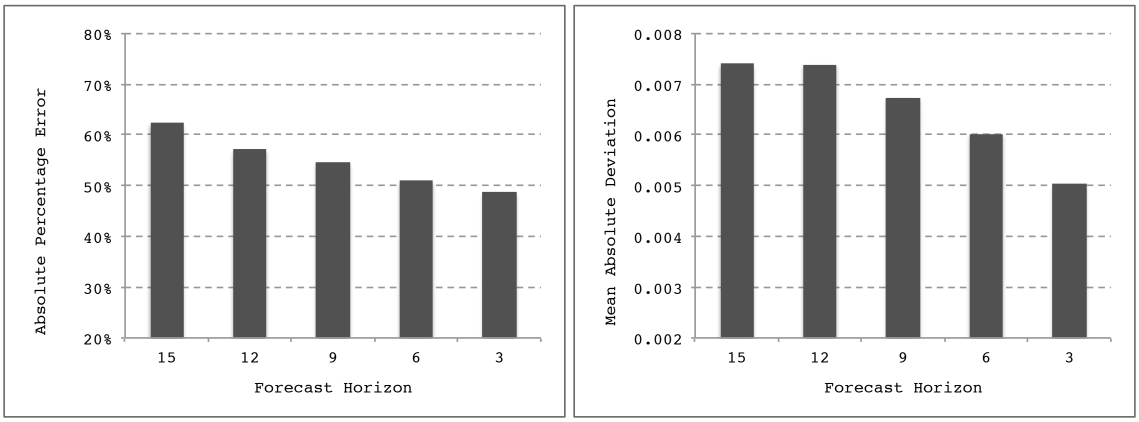

Figure 5.3.: Development of accuracy APE (on the left) and Benford MAD (on the right) in the *Large* data sample over all 5 planning horizons (from 1-3 months up to 13-15 months).

indicator and for an appropriate validation of the delivered planning data an ex-ante indicator is required. Such an ex-ante validation can be performed based on Benford's Law, if and only if, Benford's Law also reflects perceived data quality and correlates positively with data accuracy in the ideal case. Consequently, this section investigates the hypotheses **H 2.5.1**, **H 2.5.2**, and **H 2.5.3**.

In analogy to the previous section, the results for the deviation from Benford's Law MAD are separated into the development over time in Tables 5.1 and the change in dependence of the planning horizon 5.2. Moreover, the Figures 5.2 and 5.3 are not limited to the trend in accuracy values, but include the Benford trend indication as well. In Figure 5.2 it can be seen that MAD decreases from around 0.006 in 4 2008 to under 0.004 in 1 2012. This trend perfectly depicts a corresponding, highly significant decrease of $\rho = -.48$, $p < .001$ in the underlying data (cp. Table 5.1). Approximately the same decrease can be observed in Figure 5.3 for the planning horizon – more than 0.007 with 13 – 15 months planning horizon to 0.005 with 1–3 months corresponding to $\rho = .40$, $p < .001$. These results for *Large* data are supported by the other sub-samples.

We achieve strongly significant correlations for all samples, except of *CS*. Remarkably, the highest dependency can be observed in *MS* with $\rho = -.66$, $p < .001$, where we had no trend at all in accuracy. For all other samples (*Complete*/*Large*/*HC*/*DV*) the trend is more or less the same, nevertheless strong ($\rho = -.52, \rho = -.48, \rho = -.42, \rho = -.51$ and $p < .001$). Also the lowest value for small entities ($\rho = -.25$, $p < .01$) is still a medium strong and significant correlation. In contrast to the accuracy results, MAD improves in *CS* over time, too. However, the development is not significant. The analogy in the APE and MAD results is also reflected in the dependency between the time series. Despite *CS* and *MS* data, we observe a positive correlation in all samples, reaching from $\rho = .13$, $p = .126$ in *Complete* to $\rho = .47$, $p < .001$ in *HC* data. Although the dependency in *Large* and *Small* is only between small and medium strength, it is also close to significance in both cases ($\rho = .16$, $p = .075$ and $\rho = .16$, $p = .071$).

Additionally, the results in planning horizon are quite similar in APE and MAD. Again, the strongest effect can be observed in *DV* and *Large* ($\rho = .41, \rho = .40$ and $p < .001$). Moreover, *Small* and *HC* exhibit rather the same effect strength ($\rho = .38, \rho = .37$, $p < .001$), meanwhile the effect in *MS* and *Complete* is a bit smaller, nevertheless strongly significant ($\rho = .31, \rho = .30$, $p < .01$). Overall, the effect sizes are higher in MAD than in APE. Finally, like in all investigations before, *CS* data provides differing results with no effect at all ($\rho = .05$, $p < .326$).

Taking all results for the conformity with the Benford distribution into account, we can state that the perceived development over time (**H 2.5.1**) and with reduced planning horizon (**H 2.5.2**) is mostly reflected in the deviation from the expected distribution. In addition, almost all accuracy characteristics are present and there even exists a significant or at least nearly significant dependency between the two metrics (**H 2.5.3**). Consequently, they have an indication for the appropriateness of conformity with Benford's Law as an indicator for data accuracy (**RQ 2.5**). For further application of Benford's Law in corporate financial planning and for further evaluations this final result along with the divergent behaviour in *CS* data is of high importance.

Chapter 6.

Empirical Findings and Managerial Impact

The results presented in the previous chapter provide extensive insights into the characteristics of the quality metric *data accuracy*. Especially the ability of reflecting perceived data quality is shown and based upon that, the conformity with Benford's Law is introduced as an ex-ante quality metric. However, both evaluations also revealed conflicting results in minor parts of the data samples. This divergent behaviour raises the question about other, maybe external effects that might influence data quality. Consequently, the following sections present the stepwise procedure of diving deeper into the data for both samples with divergent results in Chapter 5. Starting with the investigation of external factors behind *MS* and *CS* planning accuracy in Section 6.2, Section 6.3 applies these patterns to identify the most promising data sub-sample for further investigations. Thereafter, it is of great interest to detect unknown planning patterns that allow for recommendations regarding the planning data generation to achieve business impact.

6.1. Methodology

The data foundation in this chapter mostly stays the same as in the previous chapter. However, alterations have to be made that are briefly introduced in this section. Based upon this, we present research design and investigated hypotheses in Sections 6.1.2 and 6.1.3.

6.1.1. Empirical Data Description

As above-mentioned, the available data in this chapter is similar to Chapter 5. However, the investigation of research questions **RQ 3.1** and **RQ 3.3** requires one addition and one modification:

Additional data: to decide whether planning accuracy is influenced by macro-economical insecurity or not it is necessary to generate an indicator for insecurity. In stock price theory, insecurity is often measured through volatility (Hull, 2009). Furthermore, the most important and best known stock price index is the Dow Jones that is calculated based on the stock price development of the most important American enterprises. Consequently, we adopted the Dow Jones volatility as insecurity indicator for financial planning. To achieve comparable numbers, we calculate 17 average Dow Jones volatilities per quarter based on the end of day values from first quarter 2008 to the first quarter 2012.

Modified data: to finally decide about data quality and hence to give recommendations on a decision support service, it is of particular importance to investigate the real numbers and not aggregated or relative indicators as we did in the previous chapter. Therefore, we again start with high level evaluations based on the absolute percentage error (APE), but proceed looking at the underlying exposure numbers. These exposure numbers are the actual and planned exposure E^A and E^P as they have been introduced in Section 5.1.1.

Finally, it is important to note that the investigations in this chapter are based through one single accuracy calculation incorporating only the latest planning data (short planning horizon of up to three months – cp. Sections 4.1.2 and 5.1.3). Based upon these alterations it is possible to achieve more detailed insights as they are necessary to answer the relevant research questions in this chapter.

6.1.2. Research Design

The overall research design applied in this chapter is completely the same as in the previous chapter with respect to the interferential analyses, conformity with the Benford distribution, and finally data accuracy. We again conduct the non-parametric Spearman correlation to detect dependencies, calculate MAD for Benford according to Equation 5.1 and accuracy levels as described in the Equations 5.2 to 5.4.

However, since the goal of this chapter is the extraction of individual patterns per sub-sample further analyses procedure are required. To gain such knowledge, we conduct a stepwise procedure: *(i)* the evaluation of characteristics on group level, *(ii)* the prosecution of the findings on entity level, and *(iii)* the explanation of the characteristic behaviour based on detected anomalies on the basic currency level. Section 6.1.1 already introduced special data necessary to evaluate *MS* group characteristics. Now, the investigation of seasonal effects in *CS* data requires the separation of the original time series x_i, i denoting the 17 points in time of the evaluation (Section 6.1.1), into the underlying trend or moving average MA and the seasonal effect SE

$$MA(t) = \frac{1}{4}\sum_{i=t-3}^{t} x(i), \tag{6.1}$$

$$SE(t) = x(t) - MA(t). \tag{6.2}$$

As can easily be seen in Equation 6.1, the MA can be calculated from 4 2008 ($t = 4$) onwards. This leads to a reduced time series for all evaluations based on the MA (cp. Figure 6.2). For the decisions during the steps *(ii)* and *(iii)* we make use of the extensive knowledge about subgroup and entity structures gathered within the close cooperation with our industrial partner. This knowledge resultes in the stepwise development of the expectations **E 1** to **E 4** in Section 6.1.3 and finally allows us to evaluate **H 3.2.3** and answer the corresponding research questions.

6.1.3. Hypotheses

Based on the expected patterns detected in the previous chapter, it is now possible to investigate deviations and characteristics of subgroups and specific entities. All associated hypotheses are derived in multiple steps, recursively founded on each other. Again based on expertise of knowledge workers within the holding, the expected subgroup patterns can be divided into "macro-economical patterns", "seasonal patterns", and no patterns at all. Due to the dependence on the automotive industries in Material Science we expect to detect dependencies on macro-economic insecurities. In Crop Science data we anticipate seasonal effects resulting from large deliveries in spring and payments after harvesting in fall. In Health Care (constant business, affected mainly by patent claims) and diverse entities (overlapping influences) we do not expect any patterns. Summing up, we examine the following two hypothesis for **RQ 3.1**:

H 3.1.1 – BUSINESS INSIGHTS –
Material Science planning accuracy decreases in times of macro-economical insecurity.

H 3.1.2 – BUSINESS INSIGHTS –
Crop Science planning accuracy exhibits seasonal behaviour.

However, for the decision about the further evaluation the results based on Benford's Law and weak planning efficiency are of particular importance. While conformity with Benford's Law, weak planning efficiency, and accuracy exhibit no systematic effects in *MS* data, the metrics show conflicting effects in *CS* data. The resulting contrary indication attracted our interest and led to the central research question of this chapter **RQ 3.2** (*Can conformity with Benford's Law or weak planning efficiency provide* *planning quality indications beyond data accuracy?*). Addressing this research question requires a detailed examination of systematic effects in the results. In doing so, we stepwise dive deeper into the data to identify the underlying data structures causing these effects. According to Section 6.1.1, the lowest available data level is the currency data per entity. Hence, the stepwise approach expressed in the following hypotheses terminates at this currency level:

H 3.2.1 – BUSINESS INSIGHTS –
Seasonal behaviour in Crop Science planning accuracy can be traced back to accuracy in a certain entity.

H 3.2.2 – BUSINESS INSIGHTS –
Seasonal behaviour in a certain entity's accuracy can be traced back to USD planning.

On this substantial level it is now possible to decide whether the accuracy for data quality was right or if Benford's Law and weak planning efficiency provide additional insights. Since the main contra-indication is provided for data quality in the first delivery quarter, we have to investigate the most complex and most important hypothesis:

H 3.2.3 – BUSINESS INSIGHTS –
High data accuracy in Crop Science first quarter data can be explained by delayed data adjustments.

If it is possible to accept this hypothesis, we would find an indication for answering **RQ 3.1** in the affirmative. Furthermore, the detailed investigation leading to this answer promises additional insights for **RQ 3.3**. Altogether, it would be possible to provide empirical evidence for the relevance of our quality indicators along with managerial impact.

6.2. Pattern in Subgroup Data

This section provides remarkable characteristics of data accuracy extending the findings provided in the previous chapter. Numerous external effects influence subgroup planning in dependence on the respective business line. To enable the incorporation of such effects in financial planning quality assurance, Sections 6.1 and 6.2 introduce and evaluate dependencies between different external indicators and the material and crop science subgroup.

6.2.1. Macro-Economical Insecurity influencing Material Science

As described in Section 6.1.1 we measure insecurity in macro-economy through the volatility in Dow Jones. As expected, we observe a positive correlation between the two time series. However, we even get a strong and significant positive correlation of $\rho = .52$, $p < .05$.

Considering the fact that financial planning data is generated at least one quarter in advance, it is reasonable to assume a lag of one quarter between planning insecurity and reduced planning accuracy. Hence, we shifted the accuracy time series one quarter back. This procedure leads to the two time series depicted in Figure 6.1. Therein, it can easily be seen that volatility in the Dow Jones is very closely linked to insecurity in financial planning and exhibits similar movements. Consequently, the correlation between these two time series is even greater $\rho = .62$, $p < .01$. Interestingly, a respective dependency on Benford's Law or weak planning efficiency does not exist. The correlation between Dow Jones and deviation from Benford's Law $\rho = -.08$, $p = .375$ and weak planning efficiency $\rho = .03$, $p = .452$ are both close to zero and far away from significance.

Summing up, we can accept **H 3.1.1** and for further investigations aiming at decision support for *MS* entities, it is valid to imply a dependency between macro-economical indicators and accuracy in financial planning data. Of note, however, for the success of such services it is crucial to choose the respective indicators carefully.

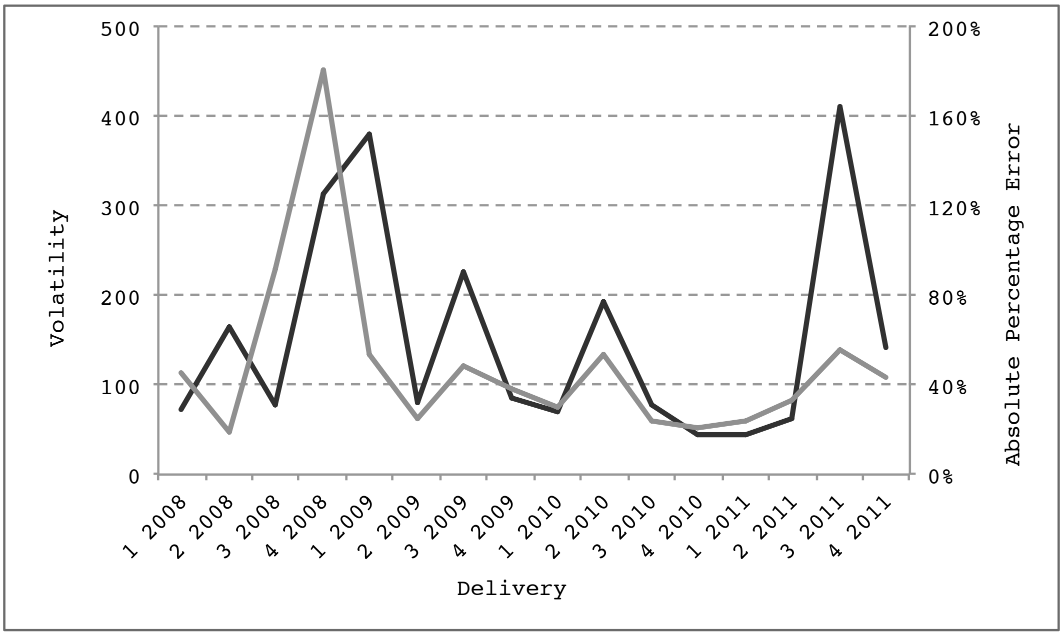

Figure 6.1.: Development of APE in *MS* planning data (light grey) compared to volatility in the Dow Jones index (dark grey) – from 1 2008 to 1 2012.

6.2.2. Seasonal Effects in Crop Science

According to Section 6.1.2, we have a clear expectation of the seasonal effects in *CS* data: the planning accuracy decreases over the year, i.e. APE increases over the year. To systematically evaluate this expectation, we calculate the influence of the progressing quarter on the complete *CS* sample, the most important entities, and on the currencies with the highest impact in entity A's data, that is the major *CS* entity. (Both classifications are based on the actual exposure volume). To visualise the seasonality of *CS* planning error, the left hand Figure in 6.2 contains three different time series: (i) the original APE in *CS* data, (ii), the trend calculated as the moving average of the original time series, and (iii) the seasonal component. As can easily be seen, the lowest error is in March deliveries meanwhile the November deliveries incorporate the largest planning error. The other two quarters range somewhere in between.In addition, the right hand Figure in 6.2 illustrates the same effects in the numbers produced by Entity A.

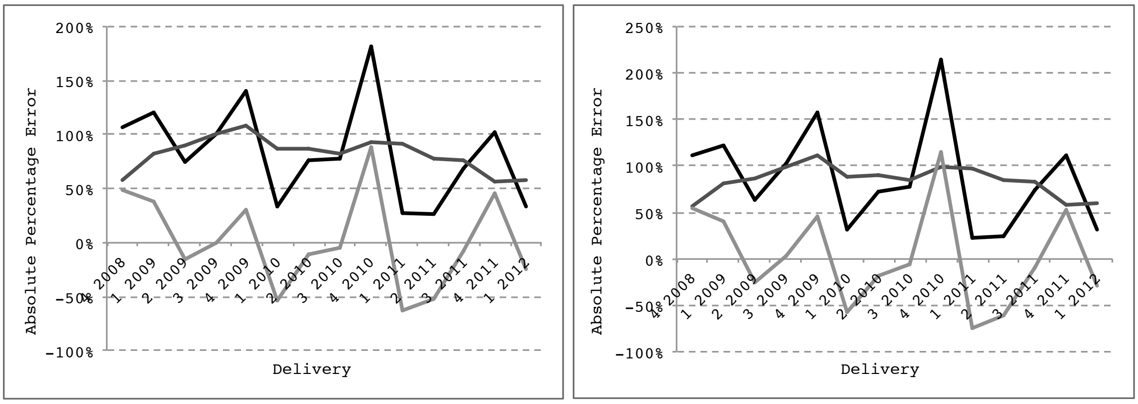

Figure 6.2.: APE (black line), trend (dark grey) and seasonal effect (light grey): Development in total *CS* planning data – on the left – compared to major *CS* company – on the right (from 4 2008 to 1 2012).

The complete dependency results listed in Table 6.1 strengthen these descriptive findings: for both, total CS data and major entity A, we observe very strong and significant seasonal increases of APE, $\rho = .70$, $p < .001$ and $\rho = .73$, $p < .001$, respectively. However, these results are not representative for all *CS* entities as easily can be seen in the results for the entities *B* and *C*: with $\rho = .24$, $p = .174$ and $\rho = -.19$, $p = .234$ the seasonal effect is barely positive or even negative. To get a differentiated view on the first entity, we calculated the same analyses for the most important currency in entity A planning to identify the key drivers for this level. The results are more heterogeneous as for the entity level: *USD* with the highest exposure is one key driver with a seasonal increase of $\rho = .72$, $p < .01$. Nevertheless, for instance, *GBP* has no significant effect ($\rho = .36$, $p = .075$).

Summing up, these results are in line with **H 3.2.1** and for further investigation it is reasonable to look for drivers of the seasonal increase in the respective entity data. Moreover, we find indication **H 3.2.2**, although, we are not able to prove that *USD* is the only driver of the seasonal increase in planning error. Nevertheless, these insights into the data dependencies along with remarkable findings regarding the relationship between Benford and weak planning efficiency on the one hand and accuracy on the other hand (cp. Section 6.3.1) enable us to start developing specific recommendations as they are introduced in the following chapter.

ρ, p	Season	Crop Science	Entity A	Entity B	Entity C	Currency USD	Currency GBP
Season	1.00	.70, .001***	.73, .000***	.24, .174	$-.19$, .234	.72, .001***	.36, .075
Crop Science		1.00	.99, .000***	.56, .010*	$-.09$, .365	.93, .000***	.38, .065
Entity A			1.00	.50, .02*	$-.16$, .268	.94, .000***	.31, .110
Entity B				1.00	.25, .164	.45, .035*	.58, .007**
Entity C					1.00	$-.28$, .141	.21, .208
Currency USD						1.00	.22, .200
Currency GBP							1.00

Table 6.1.: Cross-correlations ρ along with the respective p-values between 6 accuracy time series with 17 points in time each and a seasonal indicator ($p^* < .05$; $p^{**} < .01$; and $p^{***} < .001$).

6.3. Decision Support

The last sections provided a detailed characterization of the different data samples founded on multiple metrics. Yet, the indications are not unique for all metrics. Consequently, Section 6.3.1 presents a detailed discussion of the partly contrary indications. Thereafter, we try to generate valuable insights and to extend the previously found patterns with the goal of entity specific recommendations in Section 6.3.2 that can be incorporated in a decision support system. Additional opportunities are offered by Benford's Law in the domain of decision support services that are briefly presented in the final Section 6.3.3.

6.3.1. Contra-Indications between Benford's Law and Data Accuracy

The evaluations in the previous Section 6.2 reveal a strong seasonal increase in *CS* data and therein especially in the entity A data. During the investigation of the data characteristics it turnes out that it is a reasonable expectation that inefficiencies in the planning procedure cause these strong effects as it is expressed in **H 3.2.3**.

These expectations have their origin in contrary indications provided by Benford's Law and weak planning efficiency in contrast to accuracy. To illustrate these effects, the first line in Table 6.2 lists the correlation between progressing quarters and accuracy (APE), conformity with Benford's Law (MAD) and weak planning efficiency ($|B|$) for *CS* in total and entity A. According to Section 5.2.1 we have a strong increase in APE for both, *CS* and entity A ($\rho = .70$, $\rho = .73$, $p < .001$). For both samples MAD and $|B|$ exhibit a contrary effect or at least reveal no trend at all. MAD has a medium decrease in *CS* and entity A data ($\rho = -.22$, $p = .197$, $\rho = -.28$, $p = .141$), meanwhile $|B|$ is constant in *CS* data ($\rho = -.07$, $p = .401$) and decreases strongly and close to significance in entity A ($\rho = -.34$, $p = .09$). Summing up, accuracy indicates a decline in data quality over the year, while Benford's Law and weak planning efficiency indicate an increase in the same time interval.

This contrast is strongest in the first quarter. As can be seen in Figure 6.3 (*CS*), APE decreases strongly in the quarters *1 2010*, *1 2011*, and *1 2012* and points out an increase in accuracy. However, at the same point in time, MAD and hence the deviation from Benford's distribution increases. The same behaviour can be observed in entity A data. The resulting effect is illustrated in Figure 6.4 for weak planning efficiency that decreases in the first quarters (increase in the depicted $|B|$). These contrary indications led to the second last research question of this work **RQ 3.2** (*Can conformity with Benford's Law or weak planning efficiency provide planning quality indications beyond data accuracy?*). In this vein, the findings illustrated in Figures 6.3 and 6.4 are the starting point of the in depth evaluation in the next section and are inevitably connected to the final **RQ 3.3** (*Can entity specific recommendation be derived based on Benford's Law, weak planning efficiency, and data accuracy evaluation?*).

τ, p	Season	APE CS	MAD CS	$\|B\|$ CS	APE A	MAD A	$\|B\|$ A
Season	1.00	.70, .001***	−.22, .197	−.07, .401	.73, .000***	−.28, .141	−.34, .090
APE CS		1.00	−.40, .057	−.17, .261	.99, .000***	−.28, .139	.52, .017*
MAD CS			1.00	.23, .187	−.44, .038*	−.27, .150	.25, .172
$\|B\|$ CS				1.00	−.13, .306	−.44, .038*	.63, .003**
APE A					1.00	−.27, .150	.50, .021*
MAD A						1.00	−.06, .408
$\|B\|$ A							1.00

Table 6.2.: Cross-correlations τ along with the respective p-values between two accuracy APE, two Benford MAD, and two efficiency $|B|$ time series with 17 points in time each and a seasonal indicator ($p^{*} < .05$; $p^{**} < .01$; and $p^{***} < .001$).

Figure 6.3.: Development of accuracy APE (dark grey) and Benford MAD (light grey) in *CS* data sample over all deliveries (from 1 2008 to 1 2012).

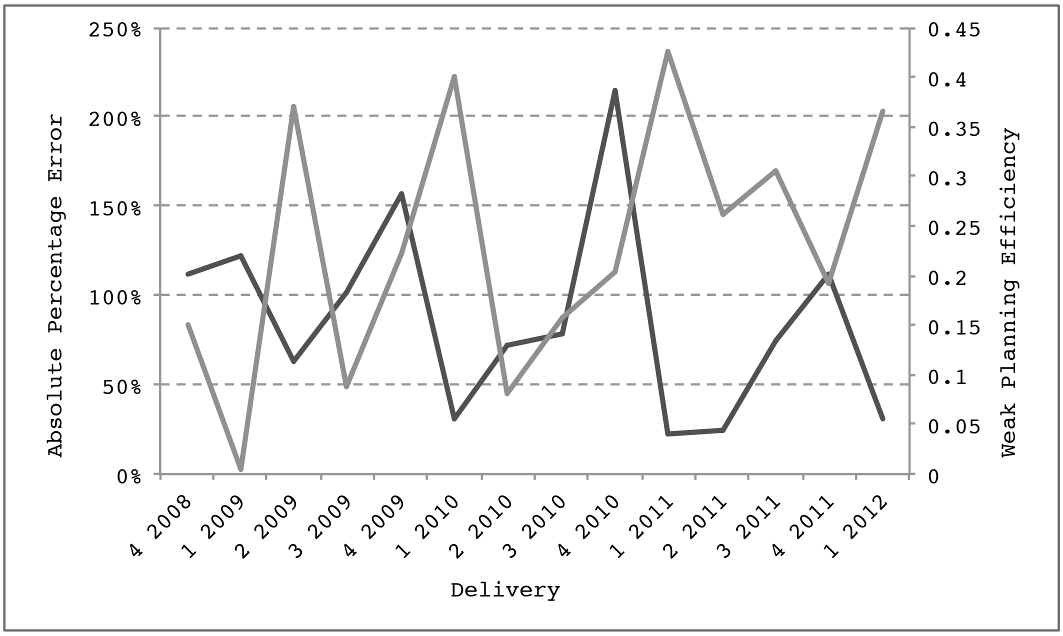

Figure 6.4.: Development of accuracy APE (dark grey) and efficiency $|B|$ (light grey) in entity *A* data sample over all deliveries (from 1 2008 to 1 2012).

6.3.2. Managerial Impact

As above-described, this section introduces a stepwise procedure leading to a specific recommendation for the data refinement of entity *A* based on indications for the confirmation of **H 3.2.3**. Although this hypothesis is a contradiction to **H 2.5.3** at first glance, both hypotheses can easily be brought into line: the general indication regarding trends evaluated in **H 2.5.3** is similar in Benford's Law and data accuracy, however, a combination of both metrics can reveal in depth results.

The procedure presented in this section can be separated into four steps, each of which is the examination of one expectation **E**. As described in detail in the previous section, the investigation is initiated by the contra-indication between the different quality metrics:

E 1 The seasonal APE increase in entity *A* data is driven by *USD* numbers.

E 2 High APE-values in the forth quarter *USD* numbers result from low actual *USD* exposure.

E 3 Systematic over-planning exists in *USD* numbers over the complete considered time period.

E 4 Planned exposure lags one quarter behind actual *USD* exposure.

The key role of *USD* in entity A data can easily be shown based on the dependencies presented in Table 6.1: the seasonal effect in entity A data ($\rho = .73$, $p < .000^{***}$) is present in *USD* data with almost the same effect size ($\rho = .72$, $p < .001^{***}$). Moreover, both time series exhibit an approximately perfect positive correlation between each other ($\rho = .94$, $p < .000^{***}$). These dependencies can be explained by the large partition of *USD* in the total entity A volume and consequently allow us to accept **E 1**.

To investigate the following expectations **E 2** and **E 3**, Figure 6.5 depicts the development of actual (E^A_{usd} – dark grey) and planned *USD* exposure (E^P_{usd} – light grey) in entity A data over all deliveries (4 2008 to 1 2012) for original data (bold line) and trend MA (dashed line), calculated as moving average (cp. Section 6.1.2). Both, E^A_{usd} and E^P_{usd}, exhibit strong seasonal behaviour. However, the expected decrease over the year expressed in **E 2** is only present in actual data ($-.55$, $p < .05$), which leads to the extremely low accuracy in the forth quarters. Moreover, in 2010, the relative difference between planned and actual MA reaches the maximum of 100% of the actual value that goes down to less than 50% at the beginning of 2012. Despite this improvement Figure 6.5 clearly shows strong over-planning of 67% on average and $E^P_{usd} < E^A_{usd}$ only holds in 1 2011 and 1 2012. These first impressions of systematic deviations between planned and actual numbers initiated further evaluations of the exposure values, expressed in **E 4**.

To prove a time lag of one quarter between E^P_{usd} and E^A_{usd}, we compare the dependency between the original time series to the dependency between the original actual time series and the planned time series moved back by one quarter. Both time series duos are presented in Figure 6.6. Especially in the latest time period starting in 2010 the time series on the right are much closer to each other. This visual difference can also be found in the statistical results: in the original time series we observe a depen-

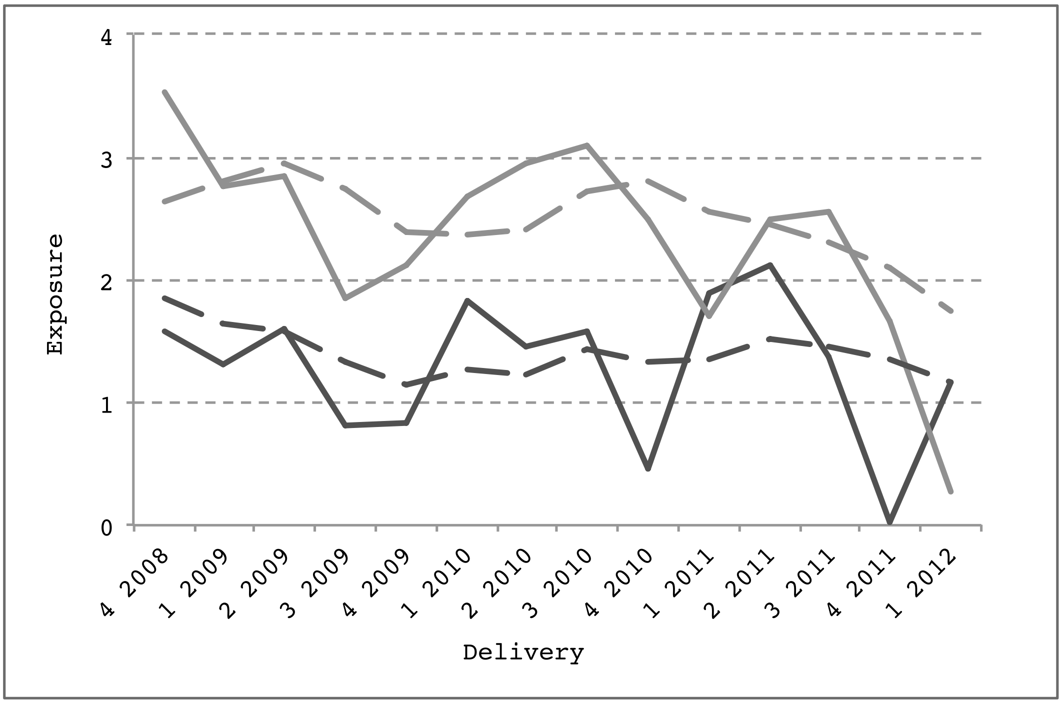

Figure 6.5.: Development of E^{A}_{usd} (dark grey) and E^{P}_{usd} (light grey) in entity A data over all deliveries (4 2008 to 1 2012) – original data (bold line) and trend MA (dashed line).

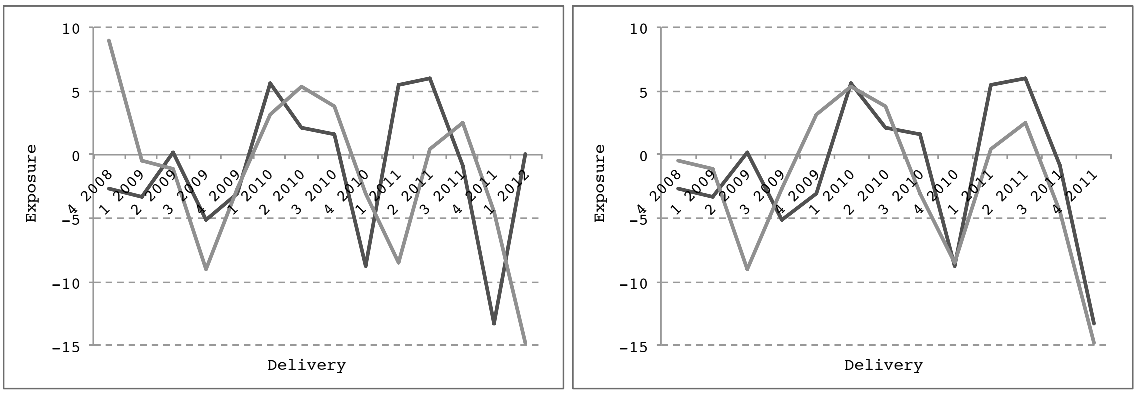

Figure 6.6.: Seasonal effect SE in E^{A}_{usd} (dark grey) and E^{usd}_{P} (light grey) in entity A data over all deliveries (4 2008 to 1 2012) – original time series (on the left) and with planned values moved one quarter back (on the right).

dency of $.46, p = .056$ and only $.35, p = .178$ (2010 to 2012). In contrast, the modified time series reveal a strong dependency of $.62, p < .05$ and even $.70, p < .05$ for the later time period (2010 to 2012) that strongly indicates the expected lag of one quarter.

Finally, summing up all results from the investigation of **E 1** to **E 4**, it seems that the strong exposure reduction in the forth quarter is not anticipated. Moreover, a faulty reduction in the first quarter along with over-planning leads to the highest accuracy in the first quarter and hence we get strong indication for confirming **H 3.2.3**. All together, these findings indicate that taking actual numbers into account does not necessarily lead to improved planning. Entity A seems to integrate actual numbers into the planning since 2010 resulting in a perfect reflection of seasonal effects with a lag of one quarter. Expert interviews revealed another possible explanation for the delay: the subsidiaries receive input updates shortly after the planing data generation. That means, the information is integrated with a gap of one quarter and this delayed integration causes the lag in the numbers. Consequently, the managerial impact can be two-fold: *(i)* it has to be checked whether the input and data generation cycle can be brought together, and *(ii)* if organizational changes are not possible, the gap can be anticipated for the final data aggregation within the holding. Independent of the realized improvement, correct integration of the seasonal effects holds significant improvement potential of 56% since 2010 (still 35% for the complete time period).

6.3.3. Benford's Law in Practice

The analyses presented in Sections 4.1 and 4.2 are the foundations for the design of our decision support service. The basic findings about Benford's Law in actual data, the increasing conformity of financial planning data to Benford's Law over the time (**H 2.1.1**), and finally the conformity with data accuracy (**H 2.5.3**) form the justification to apply digital analyses in business intelligence systems. In accordance with the discovery of characteristics for the different data sub-samples (**H 2.2.1** and **H 2.2.2**), the knowledge about the conformity of financial planning data to Benford's Law enables us to provide specified recommendations. Transferring these results into IT- based, automated decision support services, we can enrich the digit analyses with both additional statistical evaluations and expert knowledge. For example, different rounding behaviours or the consciously intended creation of duplicate numbers (e.g. through copy-and-pasting) affect the expected distribution of the delivered data. Adding expert knowledge on compliance requirements of the respective company acceptable and non-acceptable adaptations can be classified.

An exemplary, real-world decision support service taken from our industry partner could be the following: since numbers above 100,000 require further planning details due to compliance rules, numbers slightly below 100,000 may be overrepresented to save the knowledge workers time and effort. Such knowledge can be transferred and automated into the decision support service along with the knowledge on the Benford distribution in financial planning data as shown in this thesis. The manager, who accesses such a service through a business intelligence system, e.g. a corporate financial portal, may then be pointed to a following pattern: numbers beginning with a "9" (e.g. 99,000) are highly overrepresented in the first digit with a probability of 11.5% (instead of the expected 4.9%). Based upon the results for **H 2.2.1** and **H 2.2.2**, such a notification can be further enhanced, for instance, by an investigation of the second digit for positive data of a large entity. An additional application is rounding behaviour in numbers with extraordinary high volume. In a scenario with a strongly increased percentage of "0" and "5" in the second digit position, the financial planner will be pointed to all numbers with a total volume greater than, for instance, 100 million. In numbers with such a volume, rounding up or down to the above-mentioned second digits can result in a rounded volume of up to 25 million, which is not acceptable within the financial planning of our industrial partner. In both scenarios the manager is able to further investigate the issue as a result of the "alert". As expressed through the thresholds, such a decision support service can be adopted in new environments through an easy customization.

Part IV.

Finale

Chapter 7.

Conclusion & Outlook

The last years were strongly affected by the financial crisis. Insecurity and volatility in all kinds of financial markets led to a striking importance of proper liquidity and exposure planning in enterprises. The challenges for optimizing the process of data integration and validation in global companies are even greater due to distributed and heterogeneous data generation processes. To address these challenges, this work introduced a multi-stage approach for quality assurance in financial planning starting with a flexible redesign model and concluding with detailed analyses of historical and actual data.

7.1. Contribution

Data quality in financial planning has multiple dimensions and the same is true for quality assurance in financial planning. Consequently, the contribution of this work addresses *process-driven* and *data-driven* data quality separately. The first important contribution of the performed process related evaluation is the ability to actually quantify the benefit of a theoretical redesign model proposed in academia by implementing it in a real-world setting of substantial business impact. To date, such a redesign model has neither been implemented and integrated in business-relevant processes in a large company that acts worldwide nor run over a significant time producing a large set of real performance data. In that way, we are able to show that theoretical redesign models cannot only be successfully realized in highly relevant domains and enterprises but also have the potential to significantly improve *timeliness* (**RQ 1.2**), *completeness*, and *consistency* (**RQ 1.3**) in financial planning data. Consequently, we find strong indication for answering the following first research question and all corresponding sub-research questions in affirmation:

RQ 1 – INNOVATIVE BUSINESS PROCESS REDESIGN –

Can a redesign model be derived that increases data quality and similarly offers standardization and flexibility to assure practical relevance?

Thereby, redesigning financial planning processes in particular and semi-structured processes in general requires very flexible, non-standard approaches that are not provided in existing literature. This necessity led to **RQ 1.1** and the design science based development of a new redesign formalization that is applicable to all kinds of processes, independent of their structural level and their domain. Furthermore, we formulated the special redesign challenges in the six requirements in Section 3.2.1. Based on that, the model was founded on a well-structured literature representation (the objectives $\mathcal{O}$) and domain characterization (represented by the constraints $\mathcal{C}$) and a stepwise algorithm for business process redesign was introduced. The theoretical foundation for the model development were universally accepted methodologies: the design science approach of Hevner et al. (2004) and the stage activity framework by Kettinger et al. (1997). Through the fulfilment of all predefined requirements relevance and research rigour of the model were strengthened.

To evaluate the applicability, we applied the model to an example semi-structured process and documented best practices. Furthermore, a qualitative evaluation was performed through a case study at our industry partner. Therein, the development of constraints and shortcomings was documented for each single redesign step. The redesign started with the traditional financial planning process and presented a stepwise automation and implementation of the manual processes as a service-orientated architecture. As a result, the strength of the approach was shown to facilitate Financial Planning as a Service (FiPlaaS) by redesigning tasks within information systems specifically for financial planning.

In more detail, several key performance indicators of the financial planning process prior to and after the redesign were scrutinized to show and substantiate these findings. The core results of the evaluation demonstrated a reduction in Processing Time of up to 48% in working days and 19% in working weeks. Interferential analysis showed significant results for all deliveries in overall data. Moreover, we showed an improvement of the overall Planning Time of up to 19% which is also significant for three of the four regarded data samples. These changes in Planning Time are strengthened by a reduction of up to one entire working week in 80% Resolution Time.

Certainly, several business-related benefits for the industry partner come along with these results. Most importantly, the reduced Processing Time has unleashed resources that allowed additional validations to increase other financial planning data quality dimensions. The quality improvements yielded by the additional validations are strongly indicated by the significant Number of Cycles increase of 15% in November 2010 compared to November 2009. We are aware of potential biases in the data which may, for instance, be caused by organizational changes and time effects. Nevertheless, including the significantly increased Number of Cycles in November and additional expert interviews that were conducted after the study, we can demonstrate increased communication activity in September 2010 and November 2010. This activity is caused by additional intra- and inter-subsidiary as well as inter-period planning validations. Remarkably, the improvement of all time-related KPIs is strong even though the above-mentioned, new and probably time-consuming extended validations came into effect in June 2010. The reduction in process runtime unleashed valuable capacity and, hence, generates measurable business value.

Summing up, in this first part of the thesis the improvement in three of four quality dimensions (*completeness*, *consistency* and *timeliness*) through corporate financial planning redesign was shown.

For the *data-driven* quality assurance performed in this thesis, the probably most intricate contribution was to establish the conformity with Benford's Law as a quality and ex-ante accuracy indicator. Towards this goal, four interim steps had to be performed: the general conformity of financial planning data with Benford's Law had to be shown (**RQ 2.1**), the ability of indicating differences between subgroups (**RQ 2.2**). Based on this, we had to prove the representation of perceived data quality and the interdependency between data accuracy and conformity with Benford's Law (**RQ 2.4** and **RQ 2.5**). Beside all this effort, the ability of the indicator planning efficiency in detecting differences between subgroups was also investigated (**RQ 2.3**). Finally, it was possible to answer the second research question of this thesis with all associated sub-research questions in affirmation:

RQ 2 – EX-ANTE QUALITY METRIC –
Can existing quality metrics be introduced to financial planning data to generate an ex-ante quality indication?

In more detail, this work transfers analyses based on Benford's Law into a new domain: financial planning data. The results are based on a substantial set of empirical data that was provided by a globally acting, renowned large enterprise in pharmaceutical and chemical industry. Via statistical analyses of this data, we generally showed that financial planning data in fact follows Benford's Law as a contribution to the state of science. To this end, the average fulfilment rate (*AFR*) was introduced as a new quality measure to enhance the interpretation of deviations from the Benford distribution. In more detail, an increased conformity of the underlying data sample to Benford's Law over the considered time period was proven: a significant dependency between progressing time and increasing *AFR* was demonstrated for 15 of a total of 18 data treatments.

In order to transfer these findings into business relevant decision support services and to enhance the assessment of financial planning data quality, we investigated the data structure in detail in two group analyses. Through this way, analyses were conducted to validate whether decision support services that incorporate Benford's Law are appropriate to increase perceived planning data quality. In more detail, we indicated that *AFR* in the underlying financial planning data depends on the entity size and *AFR* in data delivered by large entities is higher than in the financial planning data delivered by small entities. For the second group analysis that was designed to validate whether the *AFR* in the underlying financial planning data with a positive prefix is higher than in financial planning data with negative prefix, conflicting results for first and second digit were observed. However, as expected, differences were present in the second digit (Carslaw, 1988; Thomas, 1989). Thus, the suggested separation of positive and negative numbers related to Benford's Law seems to play a significant role in the underlying domain.

In addition, weak planning efficiency was introduced as a second candidate quality measure for financial planning data. In this vein, we applied two novel measures for the analysis of rolling planning data. Yet, the gained results were very heterogeneous: negative or no efficiency development at all and few differences between subgroups. The only conformity was in *CS* data where the contra-indication between Benford and accuracy was observed. In this case, planning efficiency is completely in line with Benford's Law and provides additional motivation for the investigation beyond accuracy. Hopefully, these findings will provide a foundation for further investigations on dependencies within the financial planning process. The detection of such dependencies could leverage the implementation of new kinds of specialized decision support services in financial controlling.

For the examination of accuracy in financial planning, the planning data base was restricted to exposure data due to the limited actual data to ensure comparability. In this limited amount of data, the *AFR* was not applicable, however, this was no disadvantage as the focus was on development only. Since perceived data quality has two dimensions, we also evaluated the development of Benford conformity and accuracy in two direction: firstly, the development over time and secondly the development with decreasing planning horizon. In both ways the perceived data quality increased. In detailed analyses during this work it was shown that accuracy is in line with these expectations for five of seven sub-samples. Yet, the effect is more pronounce for the development over time, consequently, the level of significance is higher, too. Remarkably, the effects are even stronger for the conformity with Benford for all sub-samples in both directions. The only non-significant effects are present in *CS* data. Even more striking: *CS* also showed no significant trend in accuracy, for either of the two perceived data quality dimensions. This fact was the first effect pushing further evaluations towards the *CS* data sample.

The applicability of the conformity with Benford's Law as accuracy indicator was finally indicated by the dependency between the two indicators (conformity with Benford's Law and accuracy) over time: in five of seven samples the positive dependency between the indicators is significant or at least close to significance. The strength of these results is very remarkable since such systematic comparative analyses of the two indicators have not been published before. Furthermore, we can constitute a lot future work on such a set of indicators. Beside extended decision support services, systematic accuracy analyses enable additional evaluations of the redesign success.

Most important, this new quality indicator revealed empirical findings with managerial impact in the analyses of plan and actual data. Thereby, we found indication that *accuracy* as a single quality indicator is not enough to generate valuable insights into the data structure and inefficient data structures cannot be detected completely. Moreover, the introduction of the conformity with Benford's Law as additional quality indicator solved two major problems: firstly, a refined quality indication in combination with accuracy detects previously unknown inefficiencies and, secondly, a reliable ex-ante quality indication can be provided at the point of data generation. Actually, the conformity with Benford's Law is an objective quality indication in the sense that it does not depend on company standards regarding, for instance, consistency. Since data accuracy can only be calculated ex-post (actual data accessible), this additional information holds strong potential for improvement. With these significant results in mind, we can positively answer the final research question and all corresponding sub-research questions:

RQ 3 – BUSINESS INSIGHTS –
Can compositions of actual and planning data provide business insights?

As already introduced in the last paragraph, the most astonishing and as it turned out, most valuable insights during the evaluation of this superior research question were generated in answering **RQ 3.2** (*Can conformity with Benford's Law or weak planning efficiency provide planning quality indications beyond data accuracy?*). In the scenario of the most important *CS* entity *A*, the planning error significantly increased throughout the year (average increase of more than 300%), however, with a very low planning error below 50% in the first quarter. In contrast, the deviation from Benford's Law decreased over the same period of time (average decrease of more than 20%). These conflicting results initiated an in depth investigation to verify the origin of the planning error in entity *A* data with the goal of deriving an entity specific recommendation for quality improvement (**RQ 3.3**). In doing so, we detected a high dependency between accuracy in complete and *USD* data of $\rho = .94$. Moreover, accuracy in *USD* was driven by the strongly seasonal exposure development. A comparison of planned and actual exposure revealed two inefficiencies: strong over-planning and a time gap of one quarter were present in *USD* exposure. More importantly, the high accuracy in the first quarters was caused by mapping two inefficiencies in planning, a delayed exposure reduction and a simultaneous over-planning. Consequently, it did not result from accurate planning. Yet, the differences seemed to decrease in 2011. In interviews within our industry partner we figured out that at the end of 2010 a high faulty planning amount had been discovered based on manual investigations. Although, this finding already caused strong cost reduction in *USD* exposure planning, there is still much room left for further improvement: An anticipation of the seasonal effects holds an improvement potential of nearly 60%. One possible explanation for the remaining planning error is outdated input data within the data generating subsidiary. Remarkable, the delivery deadlines for financial planning at our industrial partner will be changed in the future. Although the challenges regarding timeliness of input data were known before, the findings presented in Chapter 6 had decisive influence on this development.

On the way to these interesting results, we had to conduct a large number of important analyses and research questions had to be answered. To complete the procedure behind the investigation of **RQ 3**, it was necessary to start on the highest data level and detect external factors influencing subgroup planning (**RQ 3.1**). While, *DV* and *HC* data exhibited no systematic behaviour, *MS* development was traced back to insecurity in the macroeconomy. In doing so, we showed a strongly significant dependency between the volatility in the Dow Jones index and *MS* planning accuracy of $\rho = .62$. More important for this work were the findings in *CS* data: planning accuracy revealed seasonal behaviour and decreased over the year (average increase of nearly 300%). This seasonal effect was present in entity A data with the same effect size, yet, not in all entities. The absence of this effect in the conformance with Benford's Law for *CS* and entity A data was the starting point for the very successful above-described examination.

Finally, an automated decision support will ease the detection of such faulty data in multiple ways. The research to automate the detection of promising business insights is one of the key issues of future work described in detail in the next section.

7.2. Future Work

In accordance with this entire thesis, the future work is related to the business process redesign model and the decision support based on complex data analyses. Critically assessing our approach, we were only able to implement iterations of the redesign model within one single enterprise. Nevertheless, the industry partner can be rated as archetypical for multinational enterprises and the challenges that arise from this kind of company structure. To finally evaluate the success of the redesign model, a three-fold enhancement of the presented evaluation is desirable: *(i)* the complete implementation and evaluation of all proposed services, *(ii)* the investigation of the influence of the performed changes on accuracy, and *(iii)* the roll out of the presented redesign model in additional industry environments to provide further quantitative evaluation.

As suggested by the redesign model, the implementation *(i)* is an iterative process. Therein, we will carry out a near-time implementation of the *Monitoring*, *Comment Management* and *Management Service* in general. However, it is not clear whether it will be possible to integrate the data generation process into the enterprise portal at all. More likely, the data generation will stay within the subsidiaries and the holding will provide enhanced support through the corporate financial portal. The design of evaluations and KPIs will be in analogy to Section 3.5. Furthermore, the identification of drivers behind improved data accuracy *(ii)* is crucial to design future quality assurance measures and to implement support services as presented in Section 3.3. Thereby, the assessment of accuracy improvement through the realization of an *Upload and Validation Service* requires a specific evaluation design. Through such an appropriate design biased results caused by organizational changes and additional validations, introduced continuously over time, will be avoided. Nevertheless, the realization of implementation within other industry partners might be the most challenging part. In contrast to most industry partners, the enterprise in this cooperation was not restrictive with their data and supported the research design in all phases of the project. This makes the achieved results even more interesting and valuable, however, their verification gets particularly challenging.

Equipped with detailed results of different dimensions in conformity of financial planning data with Benford's Law, we are able to address the integration of digital analyses into information systems as required by recent papers that deal with the application of Benford's Law (Nigrini, 2000; Rezaee et al., 2002). Yet, so far, digital analyses have mostly been applied to static data. With the results achieved in this work, we will be able to realize a concrete implementation of digital analyses in such a service within a business intelligence system. Beyond that, the presented results will enable a service to cope with dynamically growing data sets in the planning domain. Such a service will also take rounding amounts and contextual information like thresholds into account.

In addition to digital analyses, the decision support service will be founded on the business insights extracted from a combination of plan and actual data. That is, the previously described corrections in entity *A* data will be incorporated: either through notifications for the entity, including an explanation of over-planning and time lag, or through automated adjustment of the data on the holding side. Of course, it is highly desirable to generalize the detection of business insights. However, a holistic model requires extensive knowledge about the subsidiary characteristics. A significant step in this direction is the prototypical pattern detection presented for *CS* data in this work. Another one is initiated through the proven dependency between macroeconomy and *MS* data. To push this idea ahead, we will have to identify a set of more specific indicators that can be related to single subsidiaries. One possible indicator is the benzene price, since benzene (a crude oil constituent) is a basic component of numerous *MS* products and consequently ought to be linked to the general performance.

Beside these extensions, refinement, and implementation of procedures presented throughout this thesis, future work will focus on establishing additional quality metrics. In order to further characterize the planning behaviour within the subsidiaries, we plan to analyse the revision volume and the revision directions. Compared to weak planning efficiency, these numbers include additional information. However, they are harder to interpret, too. For illustration, Figure 7.1 visualizes four time series representing the average level of first, second, third, or fourth revisions in the complete data set over time. It shows that the behaviour during the financial crisis in 2008 significantly impacted revisions, as all corrections of planning data strongly went into the same direction. Likewise, major positive or negative peaks as can be seen in August 2007 clearly impacted the outcome of weak-form planning efficiency metrics. In this case, the peak resulted from a rather sudden M&A activity, not related to the quality of financial planning processes of involved companies. We also see that the average direction and level of revisions in a period exhibits obvious correlation, which indicates non-efficient planning processes. One lesson-learned is that such effects should be discussed with experts and removed in data-preprocessing steps prior to calculating metrics to be able to obtain proper results. Actually, data cleaning turns out not only as a crucial but also the most time-consuming and knowledge-intensive

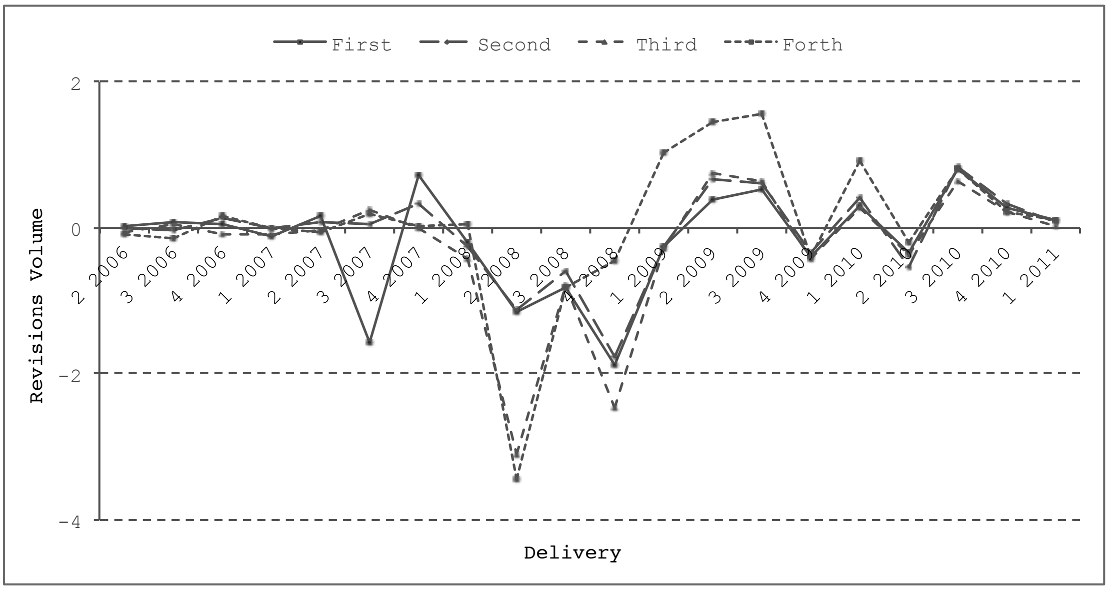

Figure 7.1.: Development of revisions i in the complete data sample over all deliveries (from 2 2006 to 1 2011).

activity. Nevertheless, an automated decision support service will tremendously decrease the complexity of planning data review and, at the same time, improve the quality of forecast data and, consequently, is of utmost importance.

The above-described complete realization of the proposed services will further increase the timeliness of the generated data and provide additional space for complex validations. One very important aspect of future work on quality improvement in financial planning data in this context is individual feedback. An inevitable step towards such an individual communication is the correct classification of entities. The accuracy calculated in this work is one metric allowing for differentiation, however, it has significant shortcomings. For instance, it does not take into account how complex the data generation itself has been. This is especially important in a very volatile business. One metric overcoming this shortcoming is *Theil's U* that compares the forecast of an entity to the dummy forecast, always taking the value of the last period:

$$U = \sqrt{\frac{\sum\limits_{t=1}^{n-1}\left(\frac{f_{t+1}-a_{t+1}}{a_t}\right)^2}{\sum\limits_{t=1}^{n-1}\left(\frac{a_t-a_{t+1}}{a_t}\right)^2}}, \tag{7.1}$$

where a_t denotes the actual value in t, a_{t+1} the actual value in $t+1$, and f_{t+1} the forecast for $t+1$. Table 7.1 compares the accuracy results to the *Theil's U* results for all four subgroup data samples, the entities *A*, *B*, and the currencies *USD*, *GBP* applied in the previous chapters. The accuracy results are identical with Table 5.1 and 5.2.

The implications of these results are very interesting in multiple ways: firstly the low values in *CS* and entity *A* imply good planning behaviour in the given volatile environment. However, currency *USD* still exhibits a higher value than the other currencies, which again strengthens the findings presented in the previous chapter. Secondly, the values > 1 for the sub-samples *HC, MS, DV* imply that the dummy forecast would have performed better in the long run. Hence, it is crucial to identify the entities and the time horizons causing the high *Theil's U* values. Based upon that, we plan to develop a clustering approach for the subsidiaries to identify the ones with great demand for planning support. Summing up, *Theil's U* provides very promising insights that should be evaluated in the future. In accordance with the findings of this thesis, it is likely that a combination of different metrics provides the most valuable insights. Moreover, we will have to discuss our findings within the industry partner to figure out which deviations can be explained and which require refinement.

One important point for the practical implications derived from the data accuracy findings in this thesis and for the comparison between subsidiaries is the pessimistic character of the calculation presented in Equations 5.2 to 5.4 and discussed in detail in Section 5.1.2. Especially since the degree to which entities are affected by negative effects resulting from internal transactions and legal requirements may vary a lot. Despite this structure of the chosen metric, today's numbers achieve an overall planning error below 40% in the first quarter 2012. This is an improvement of about 50% from the initial point of these analyses and clearly indicates a positive impact of this research project on quality assurance and a successful cooperation with the industry partner in general. Overall, the research conducted through this thesis provided valuable input towards quality assurance in financial planning. In this vein, the overall goal of research with practical impact is achieved and promising field of future work are identified.

Metric	Subgroup HC	Subgroup MS	Subgroup CS	Subgroup DV
Accuracy APE	44.0%	55.3%	84.2%	90.2%
Theil's U	1.27	1.40	0.50	1.50
	Entity A	Entity B	Currency USD	Currency GBP
Accuracy APE	103.3%	63.1%	542.2%	94.1%
Theil's U	0.43	1.21	0.79	0.43

Table 7.1.: Comparison between accuracy APE and Theil's U for eight different samples.

References

Anonymous (2010). Gaining in Stature, Feeling the Strain. *Treasury and Risk*. URL: http://www.treasuryandrisk.com/2010/01/05/gaining-in-stature,-feeling-the-strain, retrieved on 2012-08-30.

Balasubramanian, S. and M. Gupta (2005). Structural Metrics for Goal Based Business Process Design and Evaluation. *Business Process Management Journal 11*(6), 680 – 694.

Ballou, D. P. and H. L. Pazer (1985). Modeling Data and Process Quality in Multi-Input, Multi-Output Information Systems. *Management Science 31*(2), 150 – 162.

Batchelor, R. (2007). Bias in Macroeconomic Forecasts. *International Journal of Forecasting 23*(2), 189 – 203.

Batini, C., C. Cappiello, C. Francalanci, and A. Maurino (2009). Methodologies for Data Quality Assessment and Improvement. *ACM Computing Surveys (CSUR) 41*(3), 1 – 52.

Benford, F. (1938). The Law of Anomalous Numbers. *Proceedings of the American Philosophical Society 78*(4), 551 – 572.

Bhattacharyya, S., S. Jha, K. Tharakunnel, and J. Westland (2011). Data Mining for Credit Card Fraud: A Comparative Study. *Decision Support Systems 50*(3), 602 – 613.

Brenner, M. (2006). Classifying ITIL Processes; A Taxonomy under Tool Support Aspects. In *The First IEEE/IFIP International Workshop on Business-Driven IT Management (BDIM '06)*, pp. 19 – 28.

Bretschneider, S. and W. Gorr (1989). Forecasting as a Science. *International Journal of Forecasting 5*(3), 305 – 306.

Buchsein, R., H. Günther, V. Machmeier, and F. Victor (2008). *IT-Management mit ITIL ® V3 : Strategien, Kennzahlen, Umsetzung*. Wiesbaden: Vieweg+Teubner / GWV Fachverlage GmbH.

Busta, B. and R. Weinberg (1998). Using Benford's Law and Neural Networks as a Review Procedure. *Managerial Auditing Journal 13*(6), 356 – 366.

Carslaw, C. (1988). Anomalies in Income Numbers: Evidence of Goal Oriented Behavior. *The Accounting Review 63*(2), 321 – 327.

Cykana, P., A. Paul, and M. Stern (1996). DoD Guidelines on Data Quality Management. In *Proceedings of the 1996 Conference on Information Quality*, pp. 154 – 171.

Davenport, T. (1993). *Process Innovation: Reengineering Work through Information Technology*. Boston, MA, USA: Harvard Business Press.

Davenport, T., J. Harris, and S. Cantrell (2004). Enterprise Systems and Ongoing Process Change. *Business Process Management Journal 10*(1), 16 – 26.

Deiters, W. (2000). *Information Gathering and Process Modeling in a Petri Net Based Approach*, Volume 1806 of *Lecture Notes in Computer Science*, pp. 291 – 314. Berlin / Heidelberg: Springer Verlag.

Denning, P. (1997). A New Social Contract for Research. *Communications of the ACM 40*(2), 132 – 134.

Diekmann, A. (2007). Not the First Digit! Using Benford's Law to Detect Fraudulent Scientific Data. *Journal of Applied Statistics 34*(3), 321 – 329.

Drake, P. and M. Nigrini (2000). Computer Assisted Analytical Procedures Using Benford's Law. *Journal of Accounting Education 18*(2), 127 – 146.

Durtschi, C., W. Hillison, and C. Pacini (2004). The Effective Use of Benford's Law to Assist in Detecting Fraud in Accounting Data. *Journal of Forensic Accounting 5*(1), 17 – 34.

Dustdar, S. and H. Gall (2003). Architectural Concerns in Distributed and Mobile Collaborative Systems. *Journal of Systems Architecture 49*(10 – 11), 457 – 473.

Fanning, K. M. and K. O. Cogger (1998). Neural Network Detection of Management Fraud Using Published Financial Data. *International Journal of Intelligent Systems in Accounting, Finance & Management 7*(1), 21 – 41.

Feldman, M. (2000). Organizational Routines as a Source of Continuous Change. *Organization Science 11*(6), 611 – 629.

Field, A. P. (2009). *Discovering Statistics Using SPSS : (and Sex and Drugs and Rock 'n' Roll)*. Los Angeles, London: SAGE.

Fildes, R. and P. Goodwin (2008). A Few Tweaks Here and There. *Accountancy 141*(1373), 35 – 36.

Fildes, R. and H. Stekler (2002). The State of Macroeconomic Forecasting. *Journal of Macroeconomics 24*(4), 435 – 468.

Fu, T.-C. (2011). A Review on Time Series Data Mining. *Engineering Applications of Artificial Intelligence 24*(1), 164 – 181.

Fu, T.-C., F.-L. Chung, R. Luk, and C.-M. Ng (2008). Representing Financial Time Series Based on Data Point Importance. *Engineering Applications of Artificial Intelligence 21*(2), 277 – 300.

Gardyn, E. (1997). A Data Quality Handbook for a Data Warehouse. In *Proceedings of the 1997 Conference on Information Quality*, pp. 267 – 290.

Goedhart, M. H. and J. Spronk (1995). An Interactive Heuristic for Financial Planning in Decentralized Organizations. *European Journal of Operational Research 86*(1), 162 – 175.

Graham, J. and C. Harvey (2001). The Theory and Practice of Corporate Finance: Evidence from the Field. *Journal of Financial Economics 60*(2-3), 187 – 243.

Grossman, R., M. Hornick, and G. Meyer (2002). Data Mining Standards Initiative. *Communications of the ACM - Evolving Data Mining into Solutions for Insights 45*(8), 59 – 61.

Grover, V., S. Jeong, W. Kettinger, and J. Teng (1995). The Implementation of Business Process Reengineering. *Journal of Management Information Systems 12*(1), 109 – 144.

Hevner, A., S. March, J. Park, and S. Ram (2004). Design Science in Information Systems Research. *MIS Quarterly 28*(1), 75 – 105.

Hill, T. (1995). Base-Invariance Implies Benford's Law. *Proceedings of the American Mathematical Society 123*(3), 887 – 896.

Hill, T. (1998). The First Digit Phenomenon: A Century-Old Observation About an Unexpected Pattern in Many Numerical Tables Applies to the Stock Market, Census Statistics and Accounting Data. *American Scientist 86*(4), 358 – 363.

Hull, J. (2009). *Options, Futures and Other Derivatives*. Upper Saddle River, NJ: Pearson Prentice Hall.

Innig, F. (2009). IBM Finance Forum: Wie Unternehmenssteuerung agiler wird. URL: http://de.ibtimes.com/articles/8126/20090325/ibm-cognos-finance-forum-performance-management-turbulente-zeiten-david-axson-risikomanagement.htm, retrieved on 2012-08-30.

Keogh, E., K. Chakrabarti, M. Pazzani, and S. Mehrotra (2001). Dimensionality Reduction for Fast Similarity Search in Large Time Series Databases. *Knowledge and Information Systems 3*(3), 263 – 286.

Kettinger, W., J. Teng, and S. Guha (1997). Business Process Change: A Study of Methodologies, Techniques, and Tools. *MIS Quarterly 21*(1), 55 – 80.

Kim, C.-S., D. Mauer, and A. Sherman (1998). The Determinants of Corporate Liquidity: Theory and Evidence. *Journal of Financial and Quantitative Analysis 33*(3), 335 – 359.

Knight, S.-A. and J. Burn (2005). Developing a Framework for Assessing Information Quality on the World Wide Web. *Informing Science 8*, 159 – 172.

Krakar, Z. and M. Zgela (2009). Application of Benford's Law in Payment Systems Auditing. *Journal of Information and Organizational Sciences 33*(1), 39 – 51.

Kütz, M. (2011). *Kennzahlen in der IT: Werkzeuge für Controlling und Management*. Heidelberg: Dpunkt-Verlag.

Lahtela, A., M. Jantti, and J. Kaukola (2010). Implementing an ITIL-Based IT Service Management Measurement System. In *Fourth International Conference on Digital Society (ICDS '10)*, pp. 249 – 254.

Löffeler, T., R. Striemer, and W. Deiters (1998). A Framework for Identification, Classification and IT Support of Semi Structured Business Processes. *Knowledge and Process Management 5*(1), 51 – 57.

Lu, R., S. Sadiq, and G. Governatori (2008). Measurement of Compliance Distance in Business Processes. *Information Systems Management 25*(4), 344 – 355.

Lux, T. (2009). Rational Forecasts or Social Opinion Dynamics? Identification of Interaction Effects in a Business Climate Survey. *Journal of Economic Behavior & Organization 72*(2), 638 – 655.

Mandke, V. and M. Nayar (1997). Information Integrity – A Structure for its Definition. In *Proceedings of the 1997 Conference on Information Quality*, pp. 314 – 338.

Mansar, S. and H. A. Reijers (2007). Best Practices in Business Process Redesign: Use and Impact. *Business Process Management Journal 13*(2), 193 – 213.

Martin, J. (2011). Quality Assurance in Financial Planning – A Process and Data Quality Oriented Perspective. In *Proceedings of the Doctoral Consortium, Wirtschaftsinformatik 2011*, Zurich, Switzerland, pp. 130 – 139.

Martin, J., S. Betz, T. Conte, C. Gerhardt, and C. Weinhardt (2011). Objectives-Based Business Process Redesign in Financial Planning – A Case Study. In *Proceedings of the 19th European Conference on Information Systems (ECIS '11)*. Paper 24.

Martin, J. and B. Blau (2010). Dynamic Corporate Financial Plan Validation. In *Group Decision and Negotiation (GDN '10)*, pp. 216 – 220.

Martin, J., S. Caton, T. Conte, and C. Weinhardt (2011). Towards Financial Planning as a Service. In *Proceedings of the 8th International Conference on Service Computing (SCC '11)*, pp. 767 – 768.

Martin, J. and T. Conte (2012). Data Quality in Financial Planning – An Empirical Assessment Based on Benford's Law. In *Proceedings of the 20th European Conference on Information Systems (ECIS '12)*. Paper 159.

Martin, J., T. Conte, and R. Knapper (2011). Towards Objectives-Based Process Redesign. In *Proceedings of the 16th Americas Conference on Information Systems (AMCIS '11)*. Paper 150.

Martin, J., T. Conte, and A. Mazarakis (2012). Process Redesign for Liquidity Planning in Practice: An Empirical Assessment. In *Proceedings of the 24th International Conference on Advanced Information Systems Engineering (CAiSE '12)*, Volume 7328 of *Lecture Notes in Computer Science*, Gdansk, Poland, pp. 581 – 596. Springer.

Martin, J., T. Setzer, F. Teschner, T. Conte, and C. Weinhardt (2012). Decision Support Services Based on Dynamic Digital Analyses – Quality Metrics for Financial Planning Processes. In *Proceedings of the SRII Global Conference 2012*, San Jose, USA, pp. 130 – 138.

Matsumura, A. and N. Shouraboura (1996). Competing with Quality Information. In *Proceedings of the 1996 Conference on Information Quality*, pp. 72 – 86.

Newcomb, S. (1881). Note on the Frequency of Use of the Different Digits in Natural Numbers. *American Journal of Mathematics 4*(1), 39 – 40.

Niemann, F. (2009). Lösungen für das Liquiditäts-Management. URL: http://www.computerwoche.de/software/erp/1892474/, retrieved on 2012-08-30.

Nigrini, M. (1996). A Taxpayer Compliance Application of Benford's Law. *Journal of the American Taxation Association 18*(1), 72 – 91.

Nigrini, M. (2000). Continuous Auditing. *Ernst & Young Center for Auditing Research and Advanced Technology, University of Kansas*.

Nigrini, M. and L. Mittermaier (1997). The Use of Benford's Law as an Aid in Analytical Procedures. *Auditing 16*(2), 52 – 67.

Nordhaus, W. D. (1987). Forecasting Efficiency: Concepts and Applications. *The Review of Economics and Statistics 69*(4), 667 – 674.

Oberweis, A. (1996). *Modellierung und Ausführung von Workflows mit Petri-Netzen*, Volume 1 of *Teubner Reihe Wirtschaftsinformatik*. Stuttgart / Leipzig: Teubner Verlag.

O'Neill, P. and A. Sohal (1999). Business Process Reengineering – A Review of Recent Literature. *Technovation 19*(9), 571 – 581.

Pentland, B. (2003a). Conceptualizing and Measuring Variety in the Execution of Organizational Work Processes. *Management Science 49*(7), 857 – 870.

Pentland, B. (2003b). Sequential Variety in Work Processes. *Organization Science 14*(5), 528 – 540.

Phua, C., V. Lee, K. Smith, and R. Gayler (2010). A Comprehensive Survey of Data Mining-Based Fraud Detection Research. *Computer Research Repository*, 1 – 14.

Picot, A. and R. Reichwald (1984). *Bürokommunikation. Leitsätze für Anwender*. München: CW-Publikationen.

Pinkham, R. (1961). On the Distribution of First Significant Digits. *The Annals of Mathematical Statistics 32*(4), 1223 – 1230.

Pyle, D. (1999). *Data Preparation for Data Mining (The Morgan Kaufmann Series in Data Management Systems)*. Burlington, MA, USA: Morgan Kaufmann Publishers.

Raimi, R. (1976). The First Digit Problem. *American Mathematical Monthly 83*(7), 521 – 538.

Recker, J. (2010). Opportunities and Constraints: The Current Struggle with BPMN. *Business Process Management Journal 16*(1), 181 – 201.

Redman, T. (1995). Improve Data Quality for Competitive Advantage. *Sloan Management Review 36*(2), 99 – 107.

Reijers, H. and S. Limanmansar (2005). Best Practices in Business Process Redesign: An Overview and Qualitative Evaluation of Successful Redesign Heuristics. *Omega 33*(4), 283 – 306.

Rezaee, Z., A. Sharbatoghlie, R. Elam, and P. McMickle (2002). Continuous Auditing: Building Automated Auditing Capability. *Auditing 21*(1), 147 – 163.

Riege, A. (2005). Three-Dozen Knowledge-Sharing Barriers Managers Must Consider. *Journal of Knowledge Management 9*(3), 18 – 35.

Rosenkranz, C., S. Seidel, J. Mendling, M. Schäfermeyer, and J. Recker (2009). Towards a Framework for Business Process Standardization. In *Proceedings of the 5th International Workshop on Business Process Design*, pp. 53 – 63.

Sabherwal, R. and D. Robey (1993). An Empirical Taxonomy of Implementation Processes Based on Sequences of Events in Information System Development. *Organization Science 4*(4), 548 – 576.

Saroney III, E. F. (2005). Blueprint for a Successful Business. URL: http://www.nysscpa.org/cpajournal/2005/505/perspectives/p20.htm, retrieved on 2012-08-30.

Schneider-Frisse, A. (2009). Jede fünfte Firmeninsolvenz ist unnötig. *Technologie & Management – Die Verbandszeitschrift des VWI e.V.* (4), 41.

Schuh, S. (2001). An Evaluation of Recent Macroeconomic Forecast Errors. *New England Economic Review* (1), 35 – 56.

Schultz, R. L. (1992). Fundamental Aspects of Forecasting in Organizations. *International Journal of Forecasting 7*(4), 409 – 411.

Seidel, S. (2009). Toward a Theory of Managing Creativity-Intensive Processes: A Creative Industries Study. *Information Systems and E-Business Management 9*(4), 407 – 446.

Singhvi, S. S. (1972). Financial Planning in a Divisionalized Firm. *Long Range Planning 5*(4), 15–20.

Spremic, M., Z. Zmirak, and K. Kraljevic (2008). IT and Business Process Performance Management: Case Study of ITIL Implementation in Finance Service Industry. In *30th International Conference on Information Technology Interfaces (ITI '08)*, pp. 243 – 250.

Thomas, J. (1989). Unusual Patterns in Reported Earnings. *The Accounting Review 64*(4), 773 – 787.

Tsichritzis, D. (1997). *The Dynamics of Innovation*. New York, USA: Copernicus Books, Springer-Verlag.

van der Aalst, W. (2000). Workflow Verification: Finding Control – Flow Errors Using Petri-Net-Based Techniques. *Business Process Management 1806*, 19 – 128.

van der Aalst, W. and M. Weske (2001). *The P2P Approach to Interorganizational Workflows*, Volume 2068 of *Lecture Notes in Computer Science*, pp. 140 – 156. Springer Berlin / Heidelberg.

van der Aalst, W., M. Weske, and D. Grunbauer (2005). Case Handling: A New Paradigm for Business Process Support. *Data & Knowledge Engineering 53*(2), 129 – 162.

Vo, H. and H. Elsner (2007). Management of Portal Evolution – Introducing Evolution Management for the Corporate Financial Portal. *Wirtschaftsinformatik Proceedings 2007 2*, 337 – 352.

Walls, J., G. Widmeyer, and O. El Sawy (1992). Building an Information System Design Theory for Vigilant EIS. *Information Systems Research 3*(1), 36 – 59.

Wand, Y. and R. Wang (1996). Anchoring Data Quality Dimensions in Ontological Foundations. *Communications of the ACM 39*(11), 86 – 95.

Wang, R. and D. Strong (1996). Beyond Accuracy: What Data Quality Means to Data Consumers. *Journal of Management Information Systems 12*(4), 5 – 33.

Warren Liao, T. (2005). Clustering of Time Series Data – A Survey. *Pattern Recognition 38*(11), 1857 – 1874.

White, S. (2004). Introduction to BPMN. URL: http://www.omg.org/bpmn/Documents/Introduction_to_BPMN.pdf, retrieved on 2012-08-30.

Williamson, O. (1991). Comparative Economic Organization: The Analysis of Discrete Structural Alternatives. *Administrative Science Quarterly 36*(2), 269 – 296.

Witten, I., E. Frank, and M. Hall (2005). *Data Mining: Practical Machine Learning Tools and Techniques* (3rd ed.). Burlington, MA, USA: Morgan Kaufmann Publishers.

Wohed, P., W. Van der Aalst, M. Dumas, A. ter Hofstede, and N. Russell (2006). On the Suitability of BPMN for Business Process Modelling. *Business Process Management 4102*, 161 – 176.

Yelland, P. (2006). Stable Seasonal Pattern Models for Forecast Revision: A Comparative Study. *International Journal of Forecasting 22*(4), 799 – 818.

Zhang, S., C. Zhang, and Q. Yang (2003). Data Preparation for Data Mining. *Applied Artificial Intelligence 17*(5 – 6), 375 – 381.