THE ESTIMATION OF ERP LIFECYCLE COSTS

A QUANTITATIVE ANALYSIS OF COST TYPES & COST DRIVERS FOR GERMAN SMALL AND MID-SIZED COMPANIES

BENJAMIN LIEHR

Submitted in partial fulfilment of the requirements for the degree of
DOCTOR OF PHILOSOPHY

UNIVERSITY OF GLOUCESTERSHIRE
Faculty of Business, Education & Professional Studies

FEBRUARY 2017

CANDIDATE'S DECLARATION

I declare that the work in this thesis was carried out in accordance with the regulations of the University of Gloucestershire and is original except where indicated by specific reference in the text. No part of this thesis has been submitted as part of any other academic award. This thesis has not been presented to any other educational institution in the United Kingdom or overseas.

Any views expressed in this thesis are those of the author and in no way represent those of the University.

Berlin, Germany	**Signature of Candidate**
February 2017	**Benjamin Liehr**

ACKNOWLEDGEMENTS

First of all, I would like to express my sincere gratitude to all people, who supported me during the preparation of the doctoral thesis. In this context, I would like to name my supervisors, Dr. Martin Wynn and Dr. Shujun Zhang, who have supported me throughout my thesis with their patience and knowledge whilst allowing me the room to work in my own way.

In this context I would like to name the members of the University of Gloucestershire, Dr. Robin Bown and Dr. Philippa Ward, for their help in all matters of the doctoral study. The administrative staff in Bielefeld, Munich and Cheltenham always reacted quickly on queries and concerns.

I would also like to thank my family for the support they provided me through my entire life and in particular, I must acknowledge my wife, Anne, without whose support, patience, and encouragement, I would not have finished this thesis.

ABSTRACT

Contextualisation: Enterprise Resource Planning (ERP) systems have become one of the largest IT investments in recent years. Yet the implementation of this IT technology often involves some problems. Analyses and studies have identified very high cost overruns and project fiascos. There is an obvious need for better cost estimation, allowing implementing organisations a more precise or realistic specification of costs. Unfortunately, neither has a suitable model been developed nor are traditional software estimation models suitable to be transferred to ERP cost prediction. Research about this issue is relatively fragmented and the analysis of ERP costs is still in its fledging stages.

Purpose: This thesis aims to analyse the cost fields and cost drivers during the whole lifespan of an ERP lifecycle in German SMEs in the industrial sector with 30 to 1,500 employees. Different approaches for predicting ERP costs will be deduced on the basis of these findings.

Conceptual Framework: Within this thesis, the three factors "cost types", "cost drivers" and the "ERP lifecycle" are combined into one conceptual framework. The conducted systematic literature review identified the five different costs types "internal personnel costs", "external personnel costs", "hardware costs", "licence costs" and "ERP software costs" and found 35 cost drivers to be relevant in this thesis. The lifecycle is divided into three stages: evaluation, implementation and maintenance. The combination of the different cost types and different lifecycle phases results in 12 different cost fields. The cost driver candidates are analysed for their impact on each cost field.

Method: In order to access this research issue, a quantitative survey design that involved asking responsible managers in the target group about their ERP expenditures was conducted. This was accomplished by way of self-completion questionnaires provided by an online survey tool. The sampling strategy was a self-selecting one that yielded 72 eligible respondents. Based on this sample, the data was analysed for correlations, and multiple regressions were conducted using SPSS.

Findings: Firstly, this thesis identifies a cost structure of cost fields for the costs arising during each ERP lifecycle phase and for its whole lifespan. Secondly, it maps which of the identified cost drivers have an influence on each of the 12 cost fields. Thirdly, it creates a regression model of how to predict ERP costs for its whole lifespan. The developed model yields a mean magnitude of relevant error (MMRE) of 34%. Comparing this value to other approaches shows that it contributes to an improved prediction model. So far it is the best fit in ERP effort estimation.

Key Words:

- ERP
- Enterprise Resource Planning
- ERP Cost Estimation
- ERP Effort Estimation
- ERP Cost Prediction
- ERP Effort Prediction
- Cost Drivers
- Cost Types
- ERP Lifecycle
- ERP Lifespan
- ERP Cost Structure

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

DEFINITIONS

Cost Drivers	Cost drivers are the factors which cause the costs in ERP projects. Many different cost drivers are discussed in the literature. The systematic literature review found 64 cost drivers, 35 of which are pooled in this research and are the independent variables in this study.
Cost Fields	The variables emerging from the combination of cost types and lifecycle phases are named cost fields.
Cost Types	Cost types are categories into which costs are grouped. The systematic literature review found seven different cost types: internal personnel costs, external personnel costs, hardware costs, licence costs, ERP software costs, organisational change costs, and business process reengineering costs. The last two are excluded in this study. The cost types are the dependent variables in this study.
Lifecycle	ERP software usually goes through different phases in its lifetime. The literature proposes divergent suggestions regarding the classification of these phases. The existing models vary between three and six phases. This thesis regards the lifecycle of these three phases: evaluation, implementation and maintenance. The lifespan of an ERP project after its implementation is bound to be 10 years. Consequently, this thesis assumes that the maintenance phase lasts 10 years.

ABBREVIATIONS

AAD	Absolute Average Deviation
ACE	Automatic Case Selection
AD	Average Deviation
AI	Artificial Intelligence
BPR	Business Process Reengineering
CBR	Case-based Reasoning
CHF	Swiss franc
COCOTS	Constructive Commercial Off-the-Shelf Cost Model
COCOMO	Constructive Cost Model
COTS	Commercial Off-the-Shelf Model
CRM	Customer Relationship Management
DEA	Data Envelopment Analysis
DMS	Document Management System
EDI	Electronical Data Interface
ERP	Enterprise Resource Planning
H/W	Hardware
IEEE	Institute of Electrical and Electronics Engineers
IS	Information System
IT	Information Technology
LOC	Lines of Code
MdMRE	Median of the Magnitude of Relative Error
MMRE	Mean Magnitude of Relative Error
MRE	Magnitude of Relative Error
R^2	Coefficient of Determination
RQ	Research Question
SAP	Systems, Applications & Products in Data Processing
SD	Standard Deviation
SME	Small and Medium-sized Enterprises
SPSS	Statistical Package for Social Sciences
SLOC	Source Lines of Code
S/W	Software

1. INTRODUCTION

1.1 CONTEXTUALISATION

Enterprise Resource Planning (ERP) systems are one of the most important information technologies that have emerged in the previous decades (Radovilsky, 2004). They quickly developed to today's state-of-the-art technology solution (Al-Mashari, 2003; Parr & Shanks, 2000; Radovilsky, 2004) and have become one of the largest IT investments in recent years (Chung & Synder, 1999). Worldwide, a sales volume of 34.36 billion USD is expected for ERP software in the year 2017 [1].

However, the implementation of these IT systems often involves some problems. Analyses and studies have identified very high failure rates, cost overruns and project fiascos. Buckhout, Frey and Nemec (1999) for example refer to the findings of the research company Standish Group, which investigated ERP implementations in organisations having a turnover of more than 500 million USD. The study has identified an average cost overrun of 178%, an average time schedule exceeding of 230%, and an average functionality failing of 59%. Barki & Pinsonneault (2002) point to similar findings of Appleton and Zuckerman who report that more than 50% of ERP implementations fail to achieve their intended goals and 60% of the implementing organisations do not get the expected benefits.

These failures can lead to extensive financial difficulties in companies and can even result in their bankruptcy (Scott 1999; Scott & Vessey 2002, cited in Koch & Mitlöhner, 2010). The example most frequently cited in this context is probably the case of FoxMeyer Corporation. After spending more than 100 million USD over 2.5 years, the company was forced into bankruptcy and blamed the ERP implementation for its business collapse (Buckhout, et al., 1999). Unfortunately, this example is not an isolated case.

Apparently, there is a mismatch between the vendor quotation and the final costs incurred. This is a very important issue: Having little chance to predict the final costs themselves, implementing organisations rely on the calculated costs stated in the offer. How-

---

[1] According to Statistika GmbH,: http://de.statista.com/statistik/daten/studie/271721/umfrage/umsatz-mit-enterprise-resource-planning-software-weltweit/

ever, vendor invoicing is always based on the actual expenditures. Since suppliers compete against each other in order to secure an order, they may calculate their quotations optimistically, and organisations have no way of verifying these calculations.

There is obviously a need for better cost estimation that would allow implementing organisations a more precise or realistic specification of costs. In order to understand the mismatch between cost schedule and actual expense many organisations are confronted with, the already existing ERP cost estimation models have to be identified and analysed for their suitability. The current understanding of this topic is reported in the next chapter.

1.2 PREVIOUS RESEARCH

ERP cost estimation is rather an unexplored research topic and has only been scarcely discussed so far. Only 25 researchers investigated this research field between 1997 and 2010, yielding just 21 papers in this period. The approaches are varied, but none of them deliver trustworthy outcomes. Until now neither a generally accepted empirical model for estimating ERP costs nor a theoretical framework delivering reliable forecasts exists.

Research on this issue appears to be rather fragmented and just marginally developed, mainly for two reasons: The first reason is the small number of researchers and limited professional literature. There is no single book which addresses the issue. Nearly all of the few papers have been published in different journals or conference papers, indicating a lack of dominant conferences or specialised magazines. The second reason is that the professionals seem to disregard each other's studies. Most authors focus on their own research rather than taking other research into account, and this weak referencing behaviour shows.

This might be one reason why this research produces several different approach proposals which result in different outcomes.

The approaches of ERP cost estimation applied so far can be classified into two different categories: The first category transfers models that originated in other academic disciplines to ERP effort prediction, and the second category aims to design a model specifically for this issue.

The first and most important papers investigating the suitability of transferring models are the ones published by Stensrud (2001). He focused on the question of whether the existing cost prediction systems used in software development are general and flexible enough to be applied to ERP cost prediction.
Stensrud (2001) compared different parametric effort prediction systems, like linear regression, COCOMO II, COCOTS, and non-parametric effort prediction systems, like analogy and case-based reasoning (CBR), such as ANGEL, for their suitability. He concluded that most of these prediction models are not suitable for ERP effort prediction. According to Stensrud, the best results are achieved with regression analysis.
However, this opinion is controversial, and other experts like Koch & Mitlöhner (2010) claim to get better results using other methodological techniques, like analogy-based ANGEL.

The second category of approaches deals with specifically developed ERP cost estimation approaches. These can be grouped into two types: The first approach is that of theoretical suggestions, which do not yield a calculation but propose an abstract research foundation.
For example, Barki & Pinsonneault (2002) suggest a framework which puts the integration of business processes (which is the main aim of ERP in the view of the authors) in relation to their required benefits and the effort needed to achieve it. This perspective is a good contribution, but it does not provide any concrete advice regarding the question of how costs of ERP projects can be predicted. This problem also arises with an activity-based approach that aims to map the activities within an ERP project and cluster them in order to create a basis for defining the project effort.

The basic idea of the second approach is to size ERP projects by finding suitable variables which correctly map their size and complexity. Most researchers focus on testing relevant cost driver variables.
Francalanci (2001) and Equay (2008) verify potential cost driver variables by way of correlation analysis. While this provides good indications, these studies do not apply regression analysis and do allow deducing of concrete approaches in cost prediction.
The studies using regression analysis mainly verified different cost driver candidates. The only uncontroversial cost driver is the “number of users”.
Stensrud & Myrtveit (1999) identified the three variables “number of user”, “number of EDI” and “number of data conversions” to be the best variable subset in their regression

model. Widmer (2004) validated the “number of users” and “number of team members” as the two main cost drivers during an ERP lifecycle, and Koch & Mitlöhner (2010) found evidence for the cost drivers “number of users”, “type of ERP system”, “number of modules”, “number of locations” and “number of interfaces”.
Although Widmer determined that two relevant variables do not seem to indicate the whole extent of an ERP project, since this approach disregards the essential ERP attribute range of functions, his formula for calculating ERP effort yields the best results concerning the quality criteria MMRE: he attained 40%. Since he only focused on Swiss SMEs with between 25 and 100 employees, further research is needed to evaluate if these findings are generalisable with regard to bigger companies and other countries.

Together, these approaches present a very unsettled picture of research on ERP effort prediction. Although a variety of approaches was found, none of these approaches resulted in the establishment of a convincing effort estimation model. The models lack empirical evidence, and empirical research results are mainly generated through single case studies that consider only a few variables in a relatively limited period of time. A long-term and generally accepted model that can be viewed as comprehensive has not yet been developed, since all these models show some weaknesses.
Furthermore, there is a fragmented understanding of the different areas where costs are incurred within the research field of ERP cost estimation. Since most of the research focuses on specific aspects or just specific cost driver variables, an overall picture which pools the entire proposed cost driver is lacking.

Another important part of cost estimation is the point in time at which costs occur. Surprisingly, this aspect has not been given much attention in the previously developed approaches. ERP systems typically go through different lifecycle phases during their lifetime. Although literature provides differentiation in terms of the numbers used to describe the cycle, the concept of ERP lifecycles is rather similar: they all have a kind of evaluation, implementation and maintenance phase. Most authors do not regard the whole lifecycle, but focus on the implementation phase. Widmer (2004) is the only one who broadened the scope and is not limited to the implementation.
It is crucial to consider maintenance and evaluation as they are important factors for determining total costs; ignoring these factors would be fatal for ERP prediction concepts.

Since vendors have different price policies, companies can shift parts of the implementation to a later point in time when the ERP system is already in use and implementation is considered to have been completed.

In summary, this chapter shows that current ERP cost estimation approaches are unsatisfactory. A broader concept which regards all the different costs incurred with respect to all phases of the lifecycle might be a first step in achieving more precise predictions about ERP costs in the future.

1.3 RESEARCH OBJECTIVE & THEORETICAL PERSPECTIVE

The previous chapter reported a rather unsettled research field and showed that ERP cost prediction is still in the fledgling stages. Researchers need to develop a better way of predicting the costs of the whole ERP lifespan.

The aim of this research is to contribute towards a better understanding of ERP costs during its whole lifespan in German SMEs in the industrial sector with 30 to 1,500 employees. It seeks to enhance the transparency of the ERP cost structure and to improve knowledge about the factors which cause costs over the whole ERP lifespan. On the basis of these findings, approaches to predict ERP costs will be developed in this thesis.

There are several reasons for selecting the described target group. Firstly, with 18%, the German ERP market is the biggest in Europe[2].

Secondly, it can be expected that especially companies with 30 employees and more will deploy an ERP system. Smaller companies, in contrast, can be expected to rather use isolated software applications, if at all, and to not applicate integrated ERP software. The relatively high costs are the main reason why very small companies do not use ERP software. Since it was possible to consider also companies that are bigger than defined in the classical SME definition (which is up to 249 employees), it was decided to regard companies with up to 1,500 employees. This will increase the extent of the results of the thesis.

Thirdly, a special focus was placed on companies from the industrial sector because enterprises from the industrial sector benefit the most from ERP use due to the integrative

[2] http://www.erp-matchmaker.com/erp-wissen/erp-software-markt.html

character of the software and thus the integrated production support opportunities. Furthermore, especially in this sector, it can be expected that nearly every enterprise operates an ERP system[3]. Therefore, the relevance of this sector is extremely high.

Having identified correlation and regression analysis to be the most promising approach in ERP effort estimation, this thesis participates in the search of suitable or the "right" variables for mapping the size and complexity of an ERP project.
Therefore, the relevant cost drivers and their relation to ERP costs need to be determined. Both cost drivers and cost types were identified in the conducted systematic literature review.
However, costs do not only emerge during ERP implementation, but during its whole lifecycle. For that reason this research will look at the complete cycle and consider the whole lifespan.

In this thesis, these three factors, i.e. "cost types", "cost drivers" and the "ERP lifecycle", are combined into a conceptual concept.
The literature review identified five different costs types: "internal personnel costs", "external personnel costs", "hardware costs", "licence costs" and "ERP software costs". It also identified 35 cost drivers to be relevant in this thesis. The lifecycle is divided into the three stages evaluation, implementation and maintenance.

Combining the five different cost types and three different lifecycle phases will theoretically result in 15 different cost fields[4]. But the cost types hardware, software, and licence costs do not emerge during the evaluation phase; only internal and external personnel costs are bound to arise during this phase. Consequently, this thesis will only consider twelve cost fields which are graphically presented in table 1.1.

---

[3] http://www.erp-matchmaker.com/erp-wissen/erp-software-markt.html

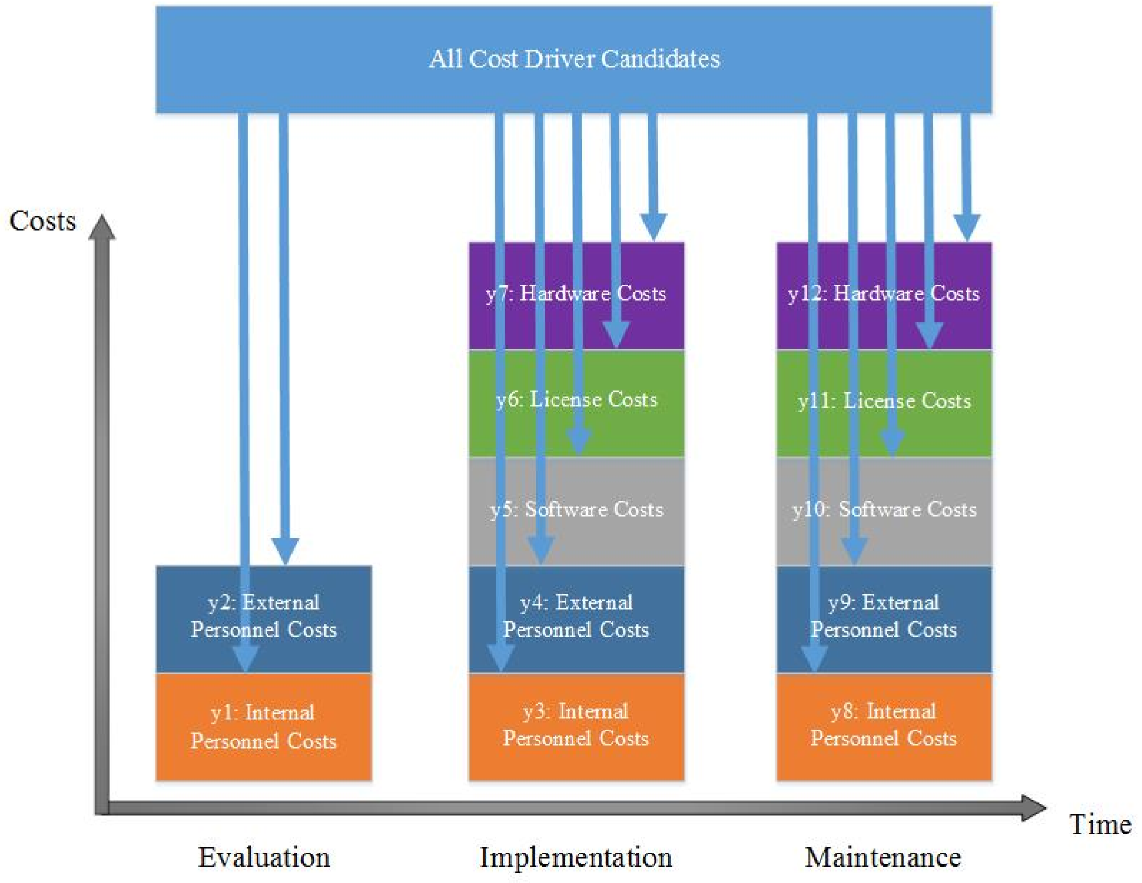
[4] The consideration of cost types within the lifecycle is referred to as "cost field" to facilitate the differentiation in the following.

Table 1.1: Cost fields of the ERP lifecycle

	Internal personnel costs	External personnel costs	ERP software costs	Licence costs	Hardware costs
Evaluation	y1	y2	-	-	-
Implementation	y3	y4	y5	y6	y7
Maintenance	y8	y9	y10	y11	y12

This breakdown will make ERP costs much more transparent as compared to just considering ERP costs as one undefined cost pool. The costs can be assigned to a certain cost field and, therewith, an ERP cost structure can be ascertained. The cost structure will show the average distribution of costs for each single cost field.
The twelve different cost fields will be analysed for their arising costs. They are thus taken as dependent variables and examined with regard to their relationships with the 35 cost drivers in a correlation analysis. Figure 1.1 shows the basic principle of this analysis.

Figure 1.1: Basic principle of research



This analysis should produce a first comprehensive and extensive overview of all emerging ERP costs. By establishing relationships between cost fields and cost drivers, ERP managers are equipped with a guideline for controlling their projects because they are aware of the factors that could influence the respective twelve cost fields.

On the basis of these findings, two approaches of cost estimation are tested. Both approaches use multiple regression models.
The first approach is to predict the costs of every single cost field. The second approach predicts the total costs of the whole ERP lifespan and distributes these costs according to the cost structure identified by answering RQ1. The result should be examined using the quality criterion MMRE. The research questions guiding this research are presented in the next chapter.

1.4 RESEARCH QUESTIONS

The aim of this research is defined on the basis of the answers to research questions. It was decided to break some questions down into smaller research sub-questions so as to receive more precise answers. The questions are presented in the following:

Table 1.2: Research question I

RQ1: What are the costs of ERP systems during their lifecycle phases?	**RQ1-1**: What are the costs for internal personnel (y1) and external personnel (y2) during the evaluation lifecycle phase?
	RQ1-2: What are the costs of internal personnel (y3), external personnel (y4), ERP software (y5), licence (y6) and hardware (y7) during the implementation lifecycle phase?
	RQ1-3: What are the costs of internal personnel (y8), external personnel (y9), ERP software (y10), licence (y11) and hardware (y12) during the ERP lifecycle maintenance phase?
	RQ 1-4: What are the costs of ERP systems during their whole lifespan?

Table 1.3: Research question II

RQ2: Which cost drivers influence ERP costs?	**RQ 2-1:** Which cost drivers influence internal personnel costs (y1) and external personnel costs (y2) in the evaluation lifecycle phase? **RQ2-2**: Which cost drivers influence internal personnel costs (y3), external personnel costs (y4), ERP software costs (y5), licence costs (y6) and hardware costs (y7) in the implementation lifecycle phase? **RQ2-3:** Which cost drivers influence internal personnel costs (y8), external personnel costs (y9), ERP software costs (y10), licence costs (y11) and hardware costs (y12) during the maintenance lifecycle phase?

Table 1.4: Research question III

RQ3: How can the identified cost drivers and cost fields be used to predict ERP costs?

An empirical quantitative research design, which is introduced in the next chapter and described in more detail in chapter 5, was selected to answer the questions.

1.5 METHODOLOGY

At the very beginning of this research, a systematic literature review was undertaken in order to get an overview of previous work done on this research issue.

The main aim of this review was to identify all relevant papers concerning this subject. Any existing cost prediction approaches, models and contributions regarding ERP systems should have been determined and analysed.

In order to get a replicable, reproducible and unbiased review, Kitchenham’s (2004) concept for undertaking a systematic review has been followed.

The systematic literature review identified different approaches in ERP cost estimation, yielded different cost types and cost drivers, and showed the disregard of the different lifecycle phases. These findings form the basis of the conceptual research framework and the research questions mentioned above.

Looking at the research questions, the nature of this research is explanatory and predictive, aiming to both test relationships between costs and/or cost types and cost drivers, and to make cost assessments based on these relationships. The applied postpositivist paradigm supports the deductive and deterministic nature of this research.

The paradigm selection often leads to a corresponding methodology. This thesis applies just quantitative research elements in order to answer the research questions.
This methodological introduction shows which methods, designs and techniques are selected. They are summarised in the list below:

- Paradigm: Postpositivist paradigm
- Research design: Quantitative approach
- Strategy of inquiry: Associational approach
- Survey design: Internet survey providing self-completion questionnaires
- Variables: Either quantified by its nature or via Likert scale
  Variables are interval or ratio scales
- Sampling strategy: Self-selecting sampling
- Data collection: EFS Survey
  Pilot Test, Pre-Survey Contact, 2 Reminders
- Data analyses: SPSS
  Coding & labelling variables
  Error reliving, avoiding third-cause fallacies
  For RQ1
  "Descriptive Statistics"
  For RQ2
  "Pearson Product Moment Correlation", "p-value"
  For RQ3
  "Multiple regression, R^2"

1.6 FINDINGS

This thesis managed to yield three main results: Firstly, it found a cost structure in RQ1. Secondly, by answering RQ2, it identified the relevant cost drivers for each cost field and thirdly, it developed a regression model for predicting ERP costs. The MMRE value of the model is 34%, which is the best identified result in ERP effort estimation so far.
The findings are based on 72 eligible respondents, which equates to a response rate of 3.8%.

In order to create a neat structure, every RQ got its own sub-chapter presenting its findings.

1.6.1 FINDINGS RQ1

RQ1 found a cost structure showing the average expenses in absolute terms and in average percentages for each cost field during each lifecycle phase. The results are presented in table 1.5.

Table 1.5 shows that the maintenance phase (a period of 10 years is considered) is generally the one in which most costs arise. It requires 53.2% of costs on average and ranges between 44.2% and 62.2% of the whole ERP lifecycle costs.
Within this phase, the licence cost is the most dominant factor with 16.3% +/-11.5. This value is followed by internal personnel costs (14.8% +15.4%/-14.8) and external personnel costs (10.9% +/- 9.5%). The lowest expenses arise for hardware (5.9% +/- 4.8%) and software (5.33% + 6.47%/-5.33) at this stage.

Table 1.5: Main finding I: Average cost structure of an ERP lifecycle

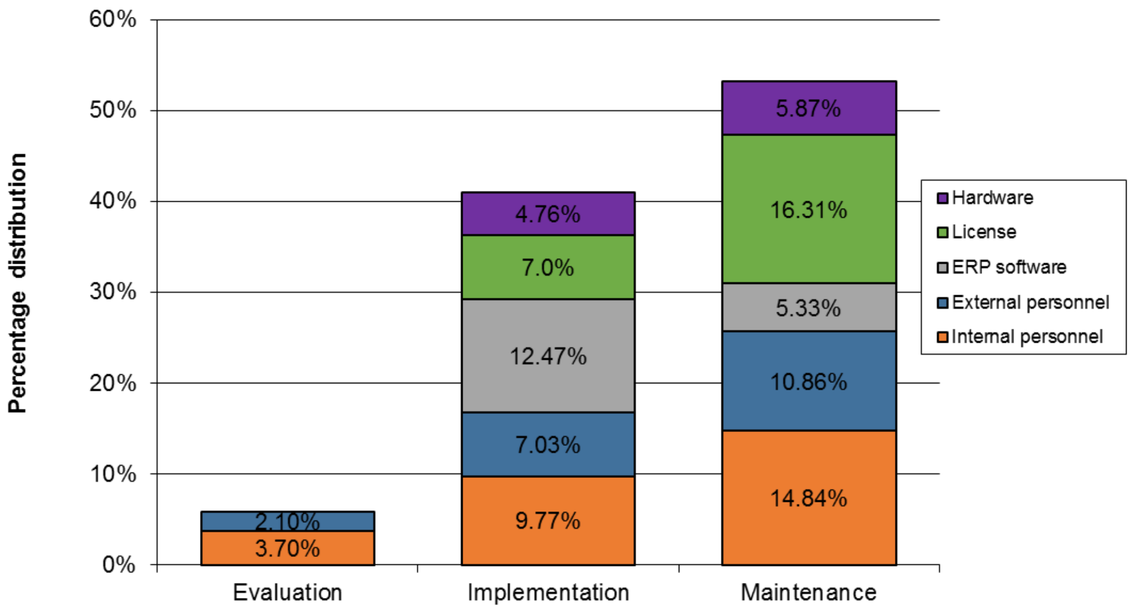
Lifecycle Phase	Cost Fields	Average Costs in absolute Terms	Average Percentage and Average Percentage Deviation	Average Percentage Deviation of each Lifecycle Phase
Evaluation Phase	y1 Internal personnel costs	35,855 EUR	3.7% ± 3.0%	5.8% ± 8.0%
	y2 External personnel costs	30,307 EUR	2.1% ± 2.6%	
Implementation Phase	y3 Internal personnel costs	136,637 EUR	9.8% ± 6.6%	41.1% ± 12.0%
	y4 External personnel costs	112,320 EUR	7.0% ± 5.0%	
	y5 ERP software costs	136,953 EUR	12.5% ± 10.1%	
	y6 Licence costs	99,454 EUR	7.0% ± 5.1%	
	y7 Hardware costs	58,250 EUR	4.8% ± 3.7%	
Maintenance Phase	y8 Internal personnel costs	390,280 EUR	14.8% ± 15.1%	53.2% ± 9.0%
	y9 External personnel costs	190,620 EUR	10.9% ± 9.5%	
	y10 ERP software costs	83,140 EUR	5.3% ± 5.9%	
	y11 Licence costs	225,550 EUR	16.3% ± 11.5%	
	y12 Hardware costs	95,830 EUR	5.9% ± 4.8%	

In total, the implementation phase costs average 40.9% +/- 12% of the whole ERP lifecycle costs. By contrast, the software costs are the cost field with the highest costs during the implementation phase. They require 12.5 +/- 10.1% on average.
This expense is followed by internal personnel costs (9.8% +/- 6.6%). Both external personnel costs and licence costs are on average 7.0%. The average deviation for external personnel costs is +/- 5.0%, and for licence costs +/- 5.1%. The lowest costs arise for hardware (4.8% +/- 3.7%) during the evaluation phase.

The lowest budget is needed for the evaluation phase. It requires from 0% to 13.8% of the total costs. With 3.7% +/- 3.0%, internal personnel costs on average require a little more of the total budget than the external personnel costs with 0% to 4.7%.

Figure 1.2 is a graphic representation of the average percentage cost structure and underlines the maintenance phase as the one with the highest expenses.

Figure 1.2: Average percentage cost distribution of the whole ERP lifecycle



These findings enable a very detailed insight into the average percentage cost distribution for the whole ERP lifecycle. The identifications create transparency about the costs arising in a time period of more than 10 years (evaluation + implementation + 10 years of maintenance) and provide further knowledge about ERP costs.

1.6.2 FINDINGS RQ2

Based on the percentage cost distribution determined with RQ1, the goal of RQ2 was to know which cost drivers are de facto responsible for the quantity of costs arising within each cost field. Thus, the cost fields were examined for their correlation with all potential cost driver candidates. In total, this thesis managed to verify 23 cost drivers showing a correlation strength of > 0.2.

The results are presented in the summary table below, which shows the correlating cost drivers and the correlation strength for each cost field.

Table 1.6: Main finding II: Correlation of cost fields and cost drivers

Cost Field	Correlating Cost Drivers	Strength of Correlation (r)
y1	x10 No. of total users	+0.408
Internal Personnel Costs	x1 No. of locations	+0.344
Evaluation Phase	x36 Satisfaction with ERP system	-0.309
	x31 Maturity of processes	-0.358
	x35 Commitment management	-0.390
	x33 Stability of organisation	-0.422
y2	x1 No. of locations	+0.406
External Personnel Costs	x10 No. of total users	+0.399
Evaluation Phase	x21 Team maturity	-0.346
	x31 Maturity of processes	-0.361
	x36 Satisfaction with ERP system	-0.426
	x33 Stability of organisation	-0.455
	x35 Commitment management	-0.460
y3	x10 No. of total users	+0.693
Internal Personnel Costs	x17 No. of external consultants	+0.643
Implementation Phase	x16 No. of internal project members	+0.526
	x2 No. of organisational units or depts.	+0.500
	x5-1 ERP system – SAP	+0.465
	x4 Revenue	+0.391
	x1 No. of locations	+0.368
	x5-6 ERP system – other	-0.334
	x33 Stability of organisation	-0.345
y4	x17 No. of external consultants	+0.719
External Personnel Costs	x10 No. of total users	+0.612
Implementation Phase	x16 No. of internal project members	+0.567
	x5-1 ERP system – SAP	+0.482
	x2 No. of organisational units or depts.	+0.443
	x12-9 Accounting module	+0.383
	x12-10 HRM module	+0.374
	x4 Revenue	+0.365
	x1 No. of locations	+0.356
	x33 Stability of organisation	-0.307
	x5-6 ERP system – other	-0.438
y5	x4 Revenue	+0.543
ERP Software Costs	x10 No. of total users	+0.536
Implementation Phase	x17 No. of external consultants	+0.383
	x2 No. of organisational units or depts.	+0.346
	x1 No. of locations	+0.342
	x11 No. of user groups	+0.322

	x16 No. of internal project members	+0.322
	x33 Stability of organisation	-0.389
y6	x17 No. of external consultants	+0.519
Licence Costs	x1 No. of locations	+0.479
Implementation Phase	x16 No. of internal project members	+0.477
	x4 Revenue	+0.464
	x10 No. of total users	+0.440
	x2 No. of organisational units or depts.	+0.314
	x12-8 Finance module	+0.313
	x31 Maturity of processes	-0.347
y7	x2 No. of organisational units or depts.	+0.751
Hardware Costs	x5-1 ERP system – SAP	+0.441
Implementation Phase	x16 No. of internal project members	+0.431
	x4 Revenue	+0.426
	x10 No. of total users	+0.415
	x17 No. of external consultants	+0.385
y8	x10 No. of total users	+0.798
Internal Personnel Costs	x2 No. of organisational units or depts.	+0.627
Maintenance Phase	x17 No. of external consultants	+0.573
	x1 No. of locations	+0.533
	x4 Revenue	+0.472
	x5-1 ERP system – SAP	+0.468
	x9 No. of EDIs	+0.332
	x16 No. of internal project members	+0.325
	x6 No. of interfaces	+0.324
	x12-12 SCM module	+0.324
	x8 No. of reports	+0.323
	x5-6 ERP system – other	-0.315
	x33 Stability of organisation	-0.361
y9	x10 No. of total users	+0.778
External Personnel Costs	x17 No. of external consultants	+0.736
Maintenance Phase	x1 No. of locations	+0.560
	x16 No. of internal project members	+0.449
	x4 Revenue	+0.397
	x2 No. of organisational units or depts.	+0.392
	x5-1 ERP system – SAP	+0.391
	x12 No. of modules	+0.371
	x12-5 DMS module	+0.323
	x35 Commitment management	-0.301
	x33 Stability of organisation	-0.333
	x5-6 ERP system – other	-0.404

y10	x17 No. of external consultants	+0.442
ERP Software Costs	x16 No. of internal project members	+0.368
Maintenance Phase	x5-1 ERP system – SAP	+0.314
y11	x10 No. of total users	+0.655
Licence Costs	x17 No. of external consultants	+0.548
Maintenance Phase	x4 Revenue	+0.522
	x1 No. of locations	+0.466
	x2 No. of organisational units or depts.	+0.365
	x16 No. of internal project members	+0.359
	x9 No. of EDIs	+0.353
	x12-5 DMS module	+0.336
	x12 No. of modules	+0.324
	x11 No. of user groups	+0.317
	x12-9 Accounting module	+0.306
y12	x17 No. of external consultants	+0.569
Hardware Costs	x16 No. of internal project members	+0.428
Maintenance Phase	x10 No. of total users	+0.405
	x4 Revenue	+0.312

As shown in the table, a total of 23 significant cost drivers were identified. Consequently, some of the discussed cost drivers can be rejected.
Of the 35 potential cost drivers discovered in the literature review, this study can eliminate 20 as not being significant.[5]

Besides identifying the relevant cost drivers influencing each cost field, the results provide a first starting point for developing a cost estimation model. The 23 relevant cost drivers might allow for a prediction of ERP costs. This approach will be specified in more detail in the next section of this chapter.

1.6.3 FINDINGS RQ3

The purpose of RQ3 was to develop a regression formula for predicting ERP costs. In accordance with the previous findings, this thesis originated two different approaches. The first approach is to estimate the costs of every single cost field, and the second ap-

[5] The two variables “x5 ERP system” and “x12 type of modules” have dummy variables. X5 has 6 and x12 has 15. The quantity of variables in total is 54.

proach is to predict the total costs of the ERP lifespan and distribute them to each cost field according to the cost structure identified in RQ1.

Both approaches were calculated by stepwise multiple regression. This method computes a prediction model by adding and/or deleting candidate predictors until one 'best' subset of predictor variables is achieved. A subset was generated for each cost field.

A comparison of the respective MMRE of the two approaches shows that it is more precise and easier to estimate the total costs of the whole ERP lifespan than to predict the costs of each cost field over the ERP lifespan.
The first approach delivers unsatisfactory results. The 12 MMREs between the estimated values and the actual values are very high. This eventuates in the disappointing realisation that the developed formulas are not suitable for accurately predicting cost fields.
Nevertheless, the deviations appear to neutralise each other and yield a good result to predict the total costs of the ERP lifecycle. The second approach generated an MMRE of 32%. Since the best identified value for ERP effort estimation so far is Widmer's result of 40%, this value seems to contribute an improvement on this research issue.
The developed formula for predicting the total costs is the following:

$$Total\ ERP\ Costs = 63{,}905 \times x_2 + 9{,}388 \times x_{10} + 82{,}041 \times x_{12} + 161{,}572 \times x_{17} - 116{,}052 \times x_{35}$$

The prediction variables are as following:
x2 No. of org. units or depts.
x10 No. of total users
x12 No. of modules
x17 No. of external consultants
x35: Commitment management
Having an adjusted R² variance of 92.2, these independent variables are able to explain 92.2% of total ERP lifecycle costs. Hence, this formula seems to be a good tool to predict the costs of an ERP lifespan.

1.7 STRUCTURE OF THESIS

This thesis is structured into seven chapters. It will firstly highlight the elementary features of ERP systems in order to provide an understanding of their basic principles.
Then the findings of the conducted systematic literature review will be presented, first looking critically at the identified approaches of ERP cost estimation and discussing their strengths and limitations. Every identified approach has its own sub-chapter so as to neatly structure this chapter. The systematic literature review identified cost types, cost drivers and lifecycle models which are reported in chapter 3.

The conceptualisation of these findings is discussed in chapter 4. It explains the research concept and provides an overview of the included cost types, cost drivers and the categorisation of the lifecycle.

To empirically validate the pooled cost types and cost drivers, chapter 5 covers the selected research methodology and design. Since all further procedures derive from the research paradigm, the section first provides a justification for the chosen paradigm. It then explains the research design and considers the quality criteria. Since this is an empirical thesis, it needs to carefully consider the treatment of the respondents, which is done in the last subsection “ethical considerations”.

Having described the methodology, this thesis then goes on to data analysis and results. It first reports on the data preparation and describes the sample, and then it answers the research questions. For a clean structure, there is a separate section for each research question. Finally, the thesis sums up the findings, discusses them, gives implications and links the findings relationship to former research, reflects on its limitations and gives recommendations in chapter 7.

2. ENTERPRISE RESOURCE PLANNING SYSTEMS

This chapter gives an overview of the basic features of ERP systems in order to provide an understanding of their specific characteristics and to outline the difference between ERP and traditional software programs, which is essential. This knowledge is important for understanding the content of this thesis.

This chapter first defines ERP and reports its basic concepts and principles, which are integration, standardisation and modular structure. Each of these will be discussed in a separate section.

2.1 DEFINITION ENTERPRISE RESOURCE PLANNING (ERP) SYSTEMS

Enterprise Resource Planning (ERP) systems are probably one of the most important information technologies that has emerged in the previous decades (Al-Mashari, 2003; Radovilsky, 2004). They became prominent in the 1990s and have quickly developed to today's state-of-the-art technology solution (Al-Mashari, 2003; Parr & Shanks, 2000; Radovilsky, 2004).

In contrast to previously available software, ERP is an integrative, process-orientated concept which is designed to connect the different activities and functions within an organisation and even throughout the entire inter-organisational supply chain in order to create integrated, efficient procedures (Buckhout, et al., 1999; Markus & Tanis, 2000; Radovilsky, 2004). In consequence, ERP is far more than an IT solution. It is rather a multidisciplinary project that "cuts into the very heart of the business, upturning policies, practices and powerbases" (Harwood, 2003, p. 1).

To understand these extensive effects, it is necessary to identify the composition of the standardised software package and modular structure of ERP.

Standardized software[6] is basically a commercial product developed to meet the generic requirements of thousands of customers rather than the specific needs of one specific organisation (Brehm, Heinzl, & Markus, 2001; Oudshoorn, 2004). The product therefore allows a certain transferability to many different organisations (Hesseler & Görtz, 2007).

With regard to the different processes within an organisation, most ERP systems are constructed modularly, which means they consist of different selectable modules (Al-Mashari, 2003; Davenport, 1998; Hesseler & Görtz, 2007). The most conventional ones cover internal processes, like manufacturing, finance, and sales, and can be expanded by inter-organisational modules, like supply chain or customer relationship management (Radovilsky, 2004). The modular structure allows organisations to install only the functions which fit their specific internal conditions and processes (Davenport, 1998).

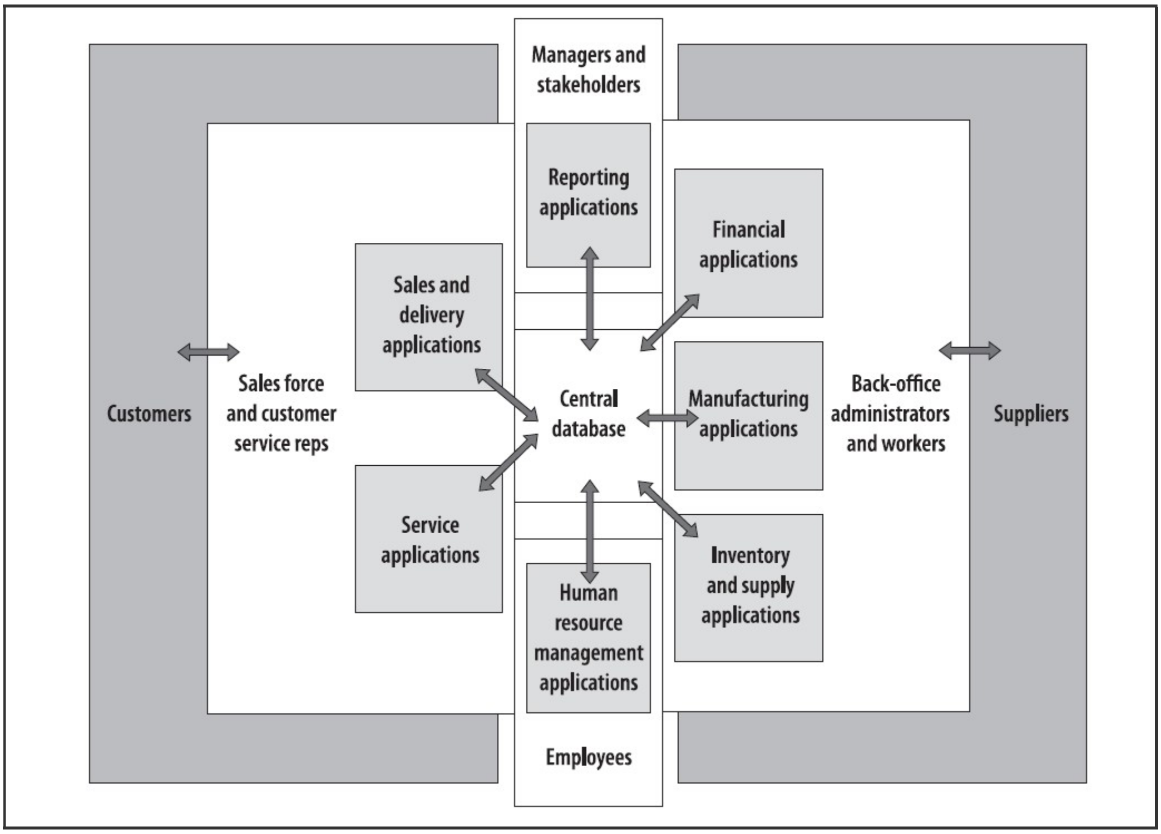
Consequently, the degree of "ERP-isation" depends on the needs of every single organisation and varies strongly from one to another. Accordingly, the ERP terminology is affected by a broad diversity of perspectives which covers the range from software to a commerce concept. Klaus, Rosemann & Gable (2000, p. 2) summarised the diverse dimensions into three aspects: "Firstly […], ERP is a commodity, a product in the form of a computer software. Second and fundamentally, ERP can be seen as a development objective of mapping all processes of data of an enterprise into a comprehensive integrative structure. Third, ERP can be seen as the key element that delivers a solution to business."

However, this thesis assumes ERP to be an "integrated computer-based system that manages internal and [where required] external organisation resources. These resources include tangible assets, financial resources, materials, and human resources. At the same time, ERP is an application and software architecture that facilitates information flows between various business functions inside and [where needed] outside an organization and, as such, is an enterprise-wide information system" (Radovilsky, 2004, p. 707). Figure 2.1 shows a generic example of the defined ERP structure.

---

[6] Standardised software is also called packaged or commercial off-the shelf (COTS) software. These terms are used interchangeably in this thesis.

Figure 2.1: Generic structure of ERP



Source: Davenport (1998) p. 125

As demonstrated in figure 2.1, ERP integrates all organisational activities and functions into a uniform system by using a centralised database and operating on a common computing platform not only within the internal business processes, but also throughout the interorganisational integration of suppliers and customers.

Such complex systems require well-designed procedures and workflows in order to succeed. They need to be elaborated carefully and defined prior to their implementation. Moreover, the standardised character of ERP often entails difficulties because the selected software package and organisational processes are not a good fit somehow. Balancing the effects of the standardised software and the special needs of an organisation is one of the most critical parts in ERP implementation and will be examined in more detail in chapter 2.2.

2.2 BASIC CONCEPTS AND PRINCIPLES

ERP has quickly grown to the state-of-the-art IT-approach all over the world (Al-Mashari, 2003; Kurbel, 2009; Parr & Shanks, 2000; Radovilsky, 2004). The background of this development is that markets and industries became more global and required organisations to adjust business and IT strategies that were congruent with these conditions (Holland, Light, & Kawalek, 1999). Moreover, this situation created more complex information, highly automated business processes and higher competition between organisations, which led to the need for more integrative IT solutions (Holland, et al., 1999; Mische, 2001; Yen, Idrus, & Yusof, 2011).

In this context, many organisations realised that their isolated IT systems provided very limited options and that, consequently, they needed to transition to newer IT strategies (Mische, 2001; Yen, et al., 2011). Up until that time, many organisations developed their software in-house. Redesigning of software in order to meet the newly manifested needs and challenges would have been very complicated, expensive and time-consuming, so that many organisations evaluated this option as inefficient (Bititci & Carrie, 1998; Holland, et al., 1999). Instead, they shifted toward standardised software packages which seemed to provide a more effective and less expensive IT approach by providing a single integrative system (Bititci & Carrie, 1998; Holland, et al., 1999).

The integrative concept, the standardised software principle and the modular structure are namely three of the most important ERP characteristics. This section will look at these three features in the following sub-sections.

2.2.1 INTEGRATION APPROACHES

The integration approach is an essential element of ERP (Davenport, 1998; Gronau, 2010). It holistically centralises all functions, processes, data and operations into one single system (Kurbel, 2009) in order to provide better, more efficient and more competitive structures (Mische, 2001). This happens not only within an organisation but also on an inter-organisational level (Radovilsky, 2004).

Systems integration is a very complex issue. It "involves a complete system of business processes, managerial practices, organisational interactions and structural alignments, and knowledge management. It is an all-inclusive process designed to create relatively seamless and highly agile processes and organisational structures that are aligned with the strategic and financial objectives of the [organisation]" (Mische, 2001, p. 5).
But the dimension of integration varies strongly from one organisation to another. Its meaning is very contextualised and an appropriate definition that fits every project and situation is hard to find (Mische, 2001).

From a technical point of view, system integration is the convention of "divergent and often incompatible technologies, applications, data, and communication into a uniform technology architecture and functional working structure" (Mische, 2001, p. 6). But as touched previously, in most cases, integration entails not only technology but involves processes, knowledge and human performance (Mische, 2001).

2.2.2 STANDARDISATION

ERP is a packaged software product which is commercially developed and distributed by different vendors. This concept revolutionised the IT market and widely replaced the previously and commonly applied in-house software[7]. Whilst custom-built software was developed to fit the special needs of the client, packaged software, like ERP, is designed to be transferable to a broad range of diverse organisations (Brehm, et al., 2001; Hesseler & Görtz, 2007).

This approach implies that different organisations must have similar workflows and business processes which can be covered by a package in a satisfactory manner. Various studies state that 80% of the requirements of organisations are met by the packaged standard software of ERP vendors (Holland, et al., 1999), which is indicative of this circumstance. New business approaches, such as just-in-time and total quality management, that emerged over the last decades were globally accepted and practiced, which might have

[7] In-house software is also called custom-built or individual software.

reinforced the idea of similarly functioning business processes (Grabot, Mayère, & Bazet, 2008; Ng, Ip, & Lee, 1999).

In order to meet the generic needs of a group of organisations, e.g. within one industry, vendors assume how most of them operate (Brehm, et al., 2001; Davenport, 1998). Their assumption is mainly based on discussions with lots of different organisations, which state their experiences, findings of researchers and academic theories. The results are crafted into "what they claim to be 'best practise'" (Al-Mashari, 2003; Bingi, Sharma, & Godla, 1999; Davenport, 1998; Glass, 1998; Markus & Tanis, 2000, p. 177). This condition is probably one of the most remarkable features of ERP, particularly in contrast to individual software: Not the customer but the vendor defines what "the best" is (Davenport, 1998, p. 4).

However, the standardised principle of ERP can induce one of the most challenging procedures in organisations (Bingi, et al., 1999; Brehm, et al., 2001; Grabot, et al., 2008; Holland, et al., 1999). When organisations work differently than their ERP system, companies need to handle the gap, often with inconvenient steep cuts (Brehm, et al., 2001). It is estimated that 20% of the processes in organisations cannot be covered by the standard version and need to be adjusted (Davis, 2005). However, adjustment is a demanding undertaking, and its degree and extent are two of the most critical factors to explain success or failure of ERP implementation (Brehm, et al., 2001; Grabot, et al., 2008; Yen, et al., 2011).

When business processes do not fit into the ERP system, organisations have two options of handling the misfit. They can either **tailor the package** to the specific business processes, which is called technical adjustment, or **re-engineer the business processes** in order to organise the work flow according to ERP, which is known as organisational adjustment (Brehm, et al., 2001; Hesseler & Görtz, 2007; Holland, et al., 1999).[8]

[8] Some authors refer to a potential third option, which is 'accepting the misfit' (Brehm, et al., 2001; Hesseler & Görtz, 2007). This option will not be considered in this thesis, because the author thinks this option induces a change which results in an organisational adjustment in order to handle the misfit somehow.

2.2.2.1 TAILORING THE ERP PACKAGE

Modification of the ERP package is a technical approach of resolving the misfit between the ERP package and the business processes. There are many suggestions for mapping the different scopes of tailoring suggested in the literature, but the most differentiated view of this issue is presented by Brehm, et al., 2001. They designed a model with nine categories which hierarchically covers the range from "small adjustments" to "major system changes", starting with "configuration" and ending with "modification".
Table 2.1 presents their nine developed tailoring levels:

Table 2.1: Tailoring types and their impact on the ERP system

Tailoring Type	Description	Examples
Configuration (Customisation or Parameterisation)	Setting of parameters (or tables) in order to choose between different executions of processes and functions in the software package	Define organisational units; create standard reports; formulate available-to-promise logic; use of a standard interface to an archive system
Bolt-ons	Implementation of third-party package designed to work with ERP system and provide industry-specific functionality	Provide ability to track inventory by product dimensions (e.g. 2 500 m lengths of cable do not equal 1 1000 m length
Screen masks	Creating new screen masks for input and output (soft copy) of data	Integrate three screens into one
Extended reporting	Programming of extended data output and reporting options	Design new report with sales revenues for specific criteria
Workflow programming	Creating non-standard workflows	Set up automated engineering-change order-specific criteria
User exits	Programming of additional software code in an open interface	Develop a statistical function for calculating particular metrics
ERP programming	Programming of additional applications without changing the source code (using the computer language of the vendor)	Create a program that calculates the phases of the moon for use in production scheduling
Interface development	Programming of interfaces to legacy systems or 3rd party products	Interface with custom-built shop-floor-system or with a CRM package
Package code modifications	Changing the source codes, ranging from small change to changing whole modules	Change error message in warning; modify production planning

Source: Brehm et al. (2001). Tailoring ERP Systems: A Spectrum of Choices and their Implications, p. 220

Tailoring might be a useful option for adjusting ERP packages to individual needs of organisations, and certain configurations are definitely inevitable. But the higher the degree of tailoring, the more disadvantages can emerge.

Tailoring has a lot of different negative aspects, like dramatically increased costs and significantly prolonged implementation times (Davenport, 1998; Somers & Nelson, 2004; Stefanou, 2001). The more tailoring is needed, “the longer it will take to roll the software out and the more it will cost to keep it up-to-date” (Bingi, et al., 1999, p. 12). In particular, tailoring might lead to difficulties with the vendor who is able to prohibit modification due to licence contracts, decline to adjust the software himself or, in cases of adjustment, refuse to provide further support for the tailored software (Brehm, et al., 2001). In the last-named case the customers are likely to have misfits with every package upgrade and need to re-implement the tailored package with every new release of the ERP software. This normally results in significantly increased maintenance cost (Glass, 1998).

In order to avoid costly and time-consuming ERP implementations as well as complicated and costly maintenance procedures, the fit between the software package processes and the way of doing business should be as close as possible.

2.2.2.2 RE-ENGINEERING THE BUSINESS PROCESS

Besides the technical tailoring, adjustments can also be made on the organisational level of an organisation. Although the idea of changing business processes in order to work with a standardised software package already emerged in the late 1980s, the extent to which it impacts organisations has dramatically increased (Holland, et al., 1999).

The adjustment of organisational processes in order to work with the ERP package is called Business Process Re-Engineering (BPR) and a key concept in ERP implementation (Schniederjans & Kim, 2003).
For defining BPR, a number of authors within the ERP literature[9] refer to the definition of Hammer & Champy (1993, p. 32) who emphasise it as “the fundamental rethinking and

9 Authors like Davenport (2000), Erkan (2009; Esteves, Pastor-Collado, & Casanovas, 2002; Ng, et al., 1999)

radical redesign of business processes to achieve dramatic improvements in critical, contemporary measures of performance, such as cost, quality, service and speed".

But, in the context of ERP implementation, BPR is controversially discussed. Whilst Al-Mashari, Al-Mudimigh & Zairi (2003, p. 362) state that benchmarked standard software is able to increase the performance of an organisation by providing "new ideas, knowledge and best practise on dealing with […] streamlining process, optimising and redesigning for more extensive benefits", Davenport emphasises that the unification of business processes, especially among competitors, can lead to the loss of unique selling points and the loss of competitiveness (Davenport, 1998).
These statements show that ERP implementation is not just a software package, but rather a powerful tool for changing business processes with sustainable impacts to the way people work (Al-Mashari, 2003; Grabot, et al., 2008; Yen, et al., 2011).

In practice, the use of any ERP imposes re-engineering of organisational processes. Since packaged software defines 'best practice' processes, organisations are forced to align their processes accordingly; thus, all organisational processes have to conform to the chosen ERP model (Bingi, et al., 1999; Glass, 1998). A lot of organisations follow ERP and do not completely reinvent their processes. The problem with BPR is that the failure rate is relatively high. It is estimated that 70% of BPR programs fail and result in strong organisational difficulties (Al-Mashari, 2003; Grabot, et al., 2008; Grant, 2002). Many organisations underestimate the time and cost increase. Especially when implementing ERP and BPR simultaneously, massive resources are put into two very successive projects at the same time (Shehab, Sharp, Supramaniam, & Spedding, 2004).

This chapter discussed the standard software package as one of the most salient ERP features and presented the two strategies of tailoring and BPR for handling this attribute. So far, BPR and tailoring were considered single strategies. In practice, they are melting policies which are combined and blended according to the specific needs of the adopting organisation. In order to avoid painful shortcomings if the selected modules do not match the business (Grabot, et al., 2008), a careful examination of the modules and their function ranges is absolutely necessary. This section will look at the modular structure of ERP in more detail in the next chapter.

2.2.3 MODULAR STRUCTURE OF ERP SYSTEMS

Most ERP systems have a modular design. This basic feature is described as modular construction system of software, in which the different functions are contained in modules and can be combined multifariously into a working overall system (Mauterer, 2002). This concept enables organisations to install only the modules and functions needed for their business processes (Davenport, 1998).

The modules and their functions differ clearly from one ERP vendor to another, which renders comparability difficult. Not only their naming, their function ranges and their working principles vary strongly (Kurbel, 2011); the selection of modules and functions is also a highly contextualised issue. Consequently, these issues complicate the selection process for organisations aiming to implement an ERP system, and a universally accepted classification is hard to find.

Authors therefore present mainly exemplary modules or functions of one specific ERP system to explain the general ERP function range. This thesis sees the most suitable classification in the work of Hesseler & Görtz (2007) who classified functions into "core functions" and "additional functions".

In their view, core functions consist of 'traditional functions' and 'inter-divisional functions'. The traditional functions include finance, logistics, production, and human resources, and the inter-divisional functions include document management, workflow management, database management, reporting, and data warehouse management.

'Additional functions' comprise e-business, customer relationship management (CRM), and supply chain management (SCM).

3. SYSTEMATIC LITERATURE REVIEW

After highlighting the elementary features of ERP systems to provide an understanding of their basic principles, the aim of this systematic literature review is to support further research on the topic of ERP cost prediction by identifying all relevant papers concerning this subject and building up the basis for the development of the conceptual framework in chapter 4.
Any existing effort prediction approaches, models and contribution for ERP systems should be identified and analysed outline which ERP effort estimation models exist and if they are suitable for accurately predicting the costs of these vast and complex systems. The following research questions should be answered within this systematic literature to meet this aim.

SLR-RQ1: How is the research environment shaped as regards ERP effort estimation?

SLR-RQ2: What ERP effort estimation models exist?

SLR-RQ3: What effort types are considered in ERP effort estimation papers?

SLR-RQ4: What are the cost drivers influencing ERP effort?

SLR-RQ5: What project or lifecycle phase is considered in ERP effort estimation papers?

In order to provide a replicable, reproducible and unbiased review, Kitchenham's (2004) concept of undertaking a systematic review in information technologies has been adapted and guides the review process and methodology, search strategy, data sources, document retrieval, the selection of studies, the including and excluding procedure, study quality assessment, data extraction & synthesis and the findings.

This chapter is structured as follows: In section 3.1 the methodology of the review process is described. This includes the presentation of the search strategy, the description of the search string and the selected data sources, the selection of studies as well as the quality assessment and data extraction.
In sections 3.2. – 3.6 the results of the 5 aforementioned research questions for the systematic literature review (SLR-RQ) are presented. Every research questions will be discussed within one single sub-section.

In section 3.7, a summative conclusion is presented enhancing the connection to this research.

3.1 METHODOLOGY AND REVIEW PROCESS

A systematic literature review aims to evaluate and interpret all available research which is relevant to a particular research question, topic area, or phenomenon of interest by using a trustworthy, rigorous, and auditable methodology (Kitchenham, 2004). Kitchenham's concept for undertaking a systematic review in information technologies has been adapted to provide a replicable, reproducible and unbiased review and is implemented in the following sub-sections.

3.1.1 SEARCH STRATEGY

The importance of the search strategy is to discover all relevant studies corresponding to the research questions without having any literature bias (Kitchenham, 2004). To ensure the inclusion of all relevant literature, a manual search was realised after having searched literature especially in different databases. A great quantity of related literature can be found with the electronic search by using an appropriate search string; it is therefore the basic approach for finding relevant studies in this review. The manual search allows for a review as to whether the authors of the literature found in databases refer to the same studies and, consequently, an examination to determine if the used search string was appropriate for finding the relevant literature. This strategy also enables the researcher to discover studies which cannot be found in the databases.

3.1.2 SEARCH STRING

In order to yield an unbiased research result, the selection of an appropriate search string used in the different databases is very important. It is necessary to use neutral, precise keywords for finding all relevant literature. The search string should include all possible synonyms and terms which can be used instead of the word the author is after. It is essential to "learn the language" by exploring the used terms, keywords, phrases and synonyms in order to conduct a search within a purposeful range. The author selects an open and unbiased wording to ensure the string does not exclude studies with "unpreferred" results. Taking into consideration these criteria, the following string has been constructed:

(ERP OR "Enterprise Resource Planning" OR "Enterprise-Resource-Planning")
AND (Cost OR Effort OR Expenditure) AND (Driver OR Predict* OR Estimat*)

The Boolean operators AND, OR, "[…]" and * were used for searching in an efficient and purposeful manner.

3.1.3 DATA SOURCES

The following databases are searched using the above-mentioned string:

- Science Direct
- ISI Web of Knowledge
- EBSCO
- Emerald Full Text
- IEEE

These databases were chosen because they are the market-leading databases. They contain a great quantity of relevant studies published in books, papers, journals, conference proceedings and magazines and therefore can provide the searcher with high-quality search results.

However, not every database allows the introduction of the search string in the above-mentioned form. In order to provide transparency, replicability and reproducibility regarding this systematic literature research, all search strings and results are documented in sufficient detail, including the unfiltered search results at this stage for a possible reanalysis.

3.1.4 DOCUMENT RETRIEVAL

The electronic search in the aforementioned databases yielded more than 392 results. These first results are defined as Step 1 in the following, meaning that the results were generated by entering the search string in the databases and searching for title or abstract.

3.1.5 SELECTION OF STUDIES

To distinguish between relevant and irrelevant results, the inclusion and exclusion criteria shown in table 3.0 are applied.

Table 3.0: Inclusion and exclusion criteria

Parameters	Inclusion Criteria	Exclusion Criteria
1. Abstract Concept	ERP stands for Enterprise Resource Planning	ERP does not stand for Enterprise Resource Planning
2. Focus	Paper is about cost estimation models or cost drivers for ERP systems	Paper is about cost estimation models or cost drivers for software development
3. Depth	Paper covers the subject of ERP cost estimation or ERP cost drivers in detail	Paper only covers the subject of ERP cost estimation or ERP cost drivers superficially or on a more general scale
4. Type of study	Primary literature	Secondary literature
5. Similarity	Paper has not been included under another title with identical content	Paper has been included under another title with identical content

Both quantitative and qualitative study types have been included in order to develop an unbiased, broad range of evidence. As the further analysis in SLR-RQ1 shows, the first relevant study was published in 1997. A time frame which defines an inclusion or exclusion of studies is not a relevant criterion for this review because consideration of the whole development from 1997 until now is important for the author to gain a broad overview.

3.1.6 PROCEDURE FOR INCLUDING AND EXCLUDING STUDIES

After 392 works have been discovered in the first step, all results now need to be reviewed in a second step to determine whether they meet the inclusion or exclusion criteria. To this end, the literature must be studied via abstract. If the abstract does not meet the inclusion criteria, the study is excluded. The full paper is studied only if the inclusion criteria are met. In this third step, papers are again excluded if the content does not correspond to the inclusion criteria. The exclusion protocol for selecting publications in this third step can be found in Appendix II. All documents are compared and existing duplicates excluded to ensure that there are no multiple publications within the selected literature.

Table 3.1: Relevant publications by data source

Name	**Date of search** *(dd/mm/yyyy)*	**Discovered (Step 1)**	**Abstract Relevant (Step 2)**	**Paper Relevant (Step 3)**	**Relevant not repeated**
ISI Web of Knowledge	06.01.2011	82	27	19	19
Business Source Complete (EBSCO)	06.01.2011	190	9	6	0
Emerald Full Text	06.01.2011	2	0	0	0
Science Direct	06.01.2011	18	1	1	0
IEEE	12.03.2011	100	14	11	0
Total		**392**	**51**	**37**	**19**

Table 3.1 shows the results of the study selection for date of search, the number of all discovered literature, the studies chosen thereof with a relevant abstract, and again the included relevant papers found in the different electronic databases which are presented separately and in total. Upon completion of the described procedure, 19 different studies were identified that directly cover the field of effort or cost prediction for ERP systems and therefore meet the inclusion criteria.

As described in section 3.1.1 “Search strategy”, above, a manual search was carried out after having read the included literature, thereby comparing the included studies with the studies the authors cite and refer to. Two new studies were obtained using this strategy; they are listed in table 3.2.

Table 3.2 Relevant studies by manual search

	ID document (ID found in Appendix 1)	New studies discovered	Relevant	Relevant not repeated
	[2]	3	1	1
	[13]	1	1	1
Total		**4**	**2**	**2**

Table 3.2 shows the results of identifying new studies by using a manual search in total numbers, the chosen studies with relevance according to the mentioned criteria, and their redundancies within the already included literature. Two of the four new studies found are not relevant because they cover business process management and software development rather than ERP implementation (exclusion criteria 1, 2 and 3). The other two studies meet the inclusion criteria and are therefore taken into consideration. Thus 21 studies are included in this literature review for answering all of the research questions.

3.1.7 STUDY QUALITY ASSESSMENT

The aim of quality assessment is to provide reliable answers for the following data synthesis and result interpreting by considering divergences in the execution of studies within design groups (Kitchenham, 2004). Different models can be used for assessing the quality of literature. Some researchers suggest weighing the quality in scores while others use checklists of factors that have to be addressed for each included publication (Kitchenham, 2004, Dixon-Woods, Agarwal, Young, Jones, & Sutton, 2004). Both methods are controversially discussed and there is no generally accepted method of how the quality of studies should be assessed (Dixon-Woods et al, 2004). Especially if qualitative and quantitative studies are included, a consensual method which accounts for all the complex details of both design approaches has not yet been developed (Dixon-Woods et al, 2004). Sandelowski, Docherty, & Emden (1997) suggest not excluding studies because of the quality in qualitative and quantitative approaches. Since there are no established criteria for assessing the quality, this approach has been followed in this literature review.

3.1.8 DATA EXTRACTION AND SYNTHESIS

In order to accurately summarise the information obtained from the included studies, the data needs to be categorised by designing an appropriate extraction form for assessing the research questions (Kitchenham, 2004). Not only characteristics for answering the research questions have been formulated, but also categories which are important for further research on the doctoral study have been created. Since numerous categories have been defined and the themes emanating from the included studies have been synthesised, different forms were designed which mostly allow a subsequent numeric evaluation of the frequency of a category. Having a numerical value is important in order to summarise the results in the research questions (Kitchenham, 2004) and merely replicate how often a characteristic or category is identified in the different included publications. For SLR-RQ1, the extraction forms are based on analysing the number of researchers and their number of published studies in years. To answer the SLR-RQ2 – SLR-RQ5, categories like approach, used methods, populations and identified cost drivers and the observation of a project phase have been designed.

3.2 SLR-RQ1: RESEARCH ENVIRONMENT OF ERP COST PREDICTION

In order to structure the first SLR research question and arrange it more neatly, it is subsectioned into three parts: "How strong is the field represented within the literature?"; "How strong is the scientific community and who are the dominant authors?" and "Where does the scientific community communicate and how is it organised?". All included papers are considered for answering the questions.

3.2.1 HOW STRONG IS THE FIELD REPRESENTED WITHIN THE LITERATURE?

The literature search conducted shows that the field of effort prediction for ERP systems is discussed scarcely among experts. As seen in section 3.1.6, only 21 relevant studies for answering all research questions were found by using the mentioned databases and via finding associated references. This quantity is interpreted as relatively poor by the author because, firstly, the search string was rather a general than a highly specialised one. Secondly, numerous companies do have cost and time overruns in ERP implementation projects (Barki & Pinsonneault, 2002). Since companies are faced with these problems worldwide, an intensive dealing with the topic within the scientific community would be obvious for the author. For that reason, 21 publications are valued as relatively moderate.

Another indication is the non-existence of books covering the implementation of ERP projects, as shown in table 3.3. Most of the professional literature can only be found in conference papers and journals.

Table 3.3: Distribution of occurrence

Occurrence	Paper references	Frequency	Proportion
Conference Papers	[2], [4], [14], [16], [21]	5	23.8%
Journal articles	[3], [5] ,[6], [8], [9], [10], [11], [12], [13], [15], [17], [18], [19], [20]	14	66.7%
Dissertations	[1],[7]	2	9.5%
Books		0	0%

3.2.2 STRENGHT OF THE SCIENTIFIC COMMUNITY AND DOMINANT AUTHORS?

To consider the professional circle, it is necessary to focus on the authors who have published relevant literature. Table 3.4 gives an overview of all authors whose studies meet the inclusion criteria, subdivided into their quantity of single-author papers, multi-author papers, the total number of publications and their paper references. The table thus shows whether a paper was written by one scientist alone or in cooperation with other scientists. The total quantity of publications was compared and taken into account in order to evaluate the dominance of an author.

Table 3.4: Most dominant researchers according to the quantity of published studies

Author	No. of single-author papers	No. of multi-author papers	Total	Paper reference
Daneva, M.	4	2	6	[9], [11], [12], [13], [14], [15]
Stensrud, E.	1	3	4	[6], [19], [20], [21]
Kusters, R. J.		3	3	[2], [4], [10]
Myrtveit, I.		3	3	[19], [20], [21]
Heemstra, F.		2	2	[4], [10]
Barki, H.		2	2	[5], [17]
Pinsonneault, A.		2	2	[5], [17]
Francalanci, C.	1		1	[3]
Hansen, T.	1		1	[16]
Arb, R. von	1		1	[1]
Widmer, T.	1		1	[7]
Janssens, G.		1	1	[10]
Equery, C.		1	1	[2]
Varone, S		1	1	[2]
Montandon, N.		1	1	[2]
Jonker, A.		1	1	[4]
Koch, S.		1	1	[8]
Mitlöhner, J.		1	1	[8]
Wieringa, R.		1	1	[13]
Wettflower, S.		1	1	[15]
Boer, S. de		1	1	[15]
Kwon, S.B		1	1	[18]
Shin, K.S.		1	1	[18]
Foss, T.		1	1	[19]
Kitchenham, B.		1	1	[19]

The table shows that the scientific community consists of 25 persons who have been occupied with the field of implementing ERP systems. Most of them have not published more than one paper. The most dominant authors are Kusters and Myrtveit who published

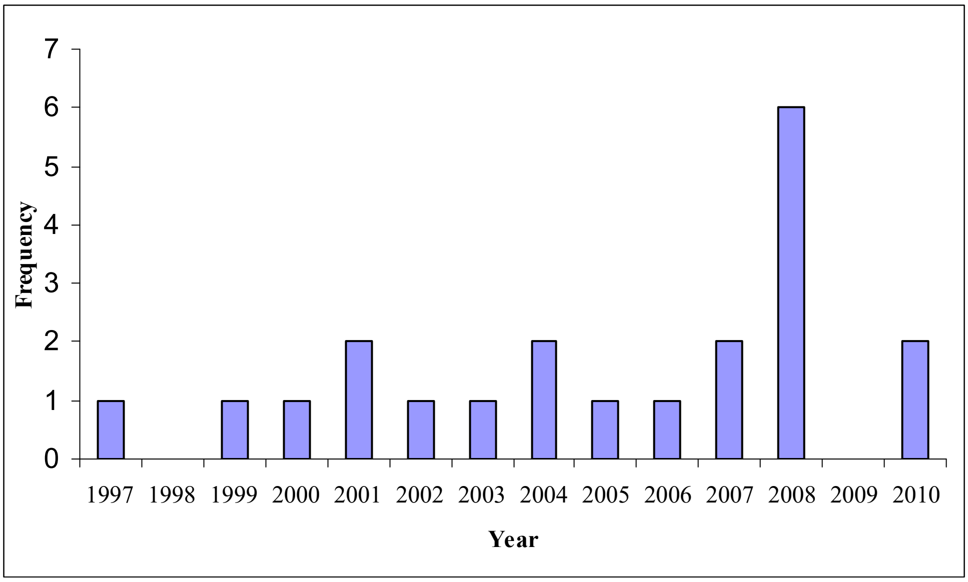
3 studies in total, followed by Stensrud who worked on 4 studies in total, and surpassed by Daneva whose total quantity is 6 studies. Since there are only 4 professionals who are publishing continuously, the author again considers the scientific community to be poorly represented.

In order to evaluate the existence of pre-eminent scientists and experts accepted by each other, it is necessary to consider if the different authors within the included literature refer to each other. Therefore the references in their studies have been compared with Appendix I.

Table 3.5 shows the paper references according to Appendix I in numbers and gives information about the number of citations in each study within the included literature and the number of authors referring to the same article, and permits a consideration of which authors are cited in whose studies.

Out of the 21 studies meeting the inclusion criteria, only 9 are cited by the other authors. 10 authors are citing Stensrud's article from 2001 [6], 6 refer to Francalanci's suggestions from 2001 [3], 5 are taking into account Daneva's study from 2007 [14], and 4 are citing the study of Daneva and Wieringa from 2008 [13] and Myrtveits's and Stensrud's literature from 1999 [20]. Paper references [6], [3] and [20] were published around the turn of the millennium. Only Daneva's [14] and Daneva's and Wieringa's [13] publications are more current from a chronological aspect. The fact that these articles are cited relatively often shows that Daneva is not only pre-eminent in this field but also that, in 5 of 6 studies, she refers to former articles written by her. For example, she cites her paper [14] in the following papers [9], [11], [12], [13] and [15]. Paper [13] is cited 3 times by her and paper 12 one time. Adjusting that fact, the author comes to the conclusion that in fact only the three relatively old studies [6], [3] and [20] do exist to which the scientific community refers.

Figure 3.0: Publication year of the included literature



Apparently, most authors focus on their own studies rather than taking into account other studies. This may be due to the fact that many different studies were written simultaneously. To review this possibility, a table was made which shows the publication year of each study. The results are shown in figure 3.0.

In general, the quantity of publications during 1997 – 2010 is constant, except for the year 2008 where 6 studies were released. Considering that 4 of these 6 articles were written by Daneva alone or by her and her co-authors, the assumption of simultaneous work on studies and consequently of not referring to each other cannot explain the fact. This leads to the question whether the authors are connected to each other or not, if they are organised in associations or organisations and if dominant journals and conferences for experts exist.

Table 3.5: Authors referring to each other in their studies

	[1]	**[2]**	**[3]**	**[4]**	**[5]**	**[6]**	**[7]**	**[8]**	**[9]**	**[10]**	**[11]**	**[12]**	**[13]**	**[14]**	**[15]**	**[16]**	**[17]**	**[18]**	**[19]**	**[20]**	**[21]**	**Total**
[1]		x		x						x												3
[2]																						0
[3]		x		x	x					x					x	x						6
[4]		x																				1
[5]																						0
[6]		x		x				x	x	x	x	x	x		x	x						10
[7]																						0
[8]																						0
[9]																						0
[10]																						0
[11]																						0
[12]															x							1
[13]								x	x		x				x							4
[14]									x		x	x	x		x							5
[15}																						0
[16]												x			x							2
[17]																						0
[18]																						0
[19]																						0
[20]						x		x											x		x	4
[21]																						0

3.2.3 WHERE DOES THE SCIENTIFIC COMMUNITY COMMUNICATE AND HOW IS IT ORGANISED ?

It might be necessary to examine this question in order to get a reference point for gaining further studies, information and an access to the research community. Therefore, this sub-section should point out the existence of relevant associations, journals and conferences in which the ERP effort prediction community is organised. Often relevant associations organise conferences and publish their own magazines and journals. Having not identified any associations or other societies and unions focusing on ERP effort prediction, the existence of dominant journals and conferences might refer to a society. Therefore, the journals and conference proceedings in which the included literature was published is analysed exactly.

The occurrence of the different articles published in journals is shown in table 3.6. It can be seen that all studies were mainly printed in different journals, except for the journal "International Journal of Enterprise Information Systems" where two of the included studies were released.

Table 3.6: Existence of dominant journals

Name of Journal	Paper Reference	Frequency
International Journal of Enterprise Information Systems	[8], [10]	2
Journal of Information Technology	[3]	1
Computer and Information Science	[5]	1
Information and Software Technology	[6]	1
Journal of Software Maintenance and Evolution: Research and Practice	[9]	1
Requirements engineering: Foundation for software quality	[11]	1
Product-focused software process improvement	[12]	1
Software Quality Journal	[13]	1
Software process and product measurement	[15]	1
Organization Science	[17]	1
Artificial intelligence and simulation	[18]	1
Empirical software engineering	[19]	1
IEEE Transactions on Software Engineering	[20]	1
Total		**14**

To consider whether dominant conferences exist in the field of ERP implementation, the relevant literature published in conference proceedings has to be taken into account. Table 3.7 gives an overview of the included studies released in conference proceedings.

Table 3.7: Existence of dominant conferences

Conference / Proceedings	Paper Reference	Frequency
International Conference on Enterprise Information Systems	[2], [4]	2
International Symposium on Empirical Software Engineering and Measurement	[14]	1
On the Move to Meaningful Internet Systems	[16]	1
International Software Metrics Symposium	[21]	1
Total		**5**

The result is similar to the result of table 3.6. Indeed, the field of ERP implementation was discussed twice at the "International Conference on Enterprise Information Systems" but all other conference papers were lectured during diverse conferences and, as a consequence, published in different conference proceedings. A clear dominance of one conference is therefore not recognisable.

To summarise the research environment: 25 persons had been occupied with the field of ERP effort estimation from 1997 – 2010. Most of them published not more than one study during this time. Therefore, four researchers are considered to be dominant having published 3, 4 and 6 studies. In considering whether pre-eminent scientists exist, the referring behaviour of all included authors has been studied. Except for three articles written around the turn of the millennium, the different authors sparsely cite each other.

The "scientific community" is not organised in associations or societies which are occupied with ERP systems; it is not equipped with dominant journals, and dominant conferences do not clearly exist. The author evaluates these facts as being relatively poor and comes to the conclusion that a scientific community exists only marginally.

3.3 SLR-RQ2: WHAT ERP EFFORT ESTIMATION MODELS EXIST?

The systematic literature review identified just 14 papers dealing with ERP cost estimation models. In total, it found seven different approaches to ERP effort estimation, which are listed below:

- Transferability of estimation models for custom-built software (COCOMO II and COCOTS)
- Case-based Reasoning (analogy-based models)
- Social Choice Theory
- Data Envelopment Analysis
- Activity-based ERP Effort Estimation
- Organisational Integration
- Correlation and Regression Analysis

Table 3.8 gives an overview of the models, their researchers and the frequency of their appearance in the literature.

Independent of its specific approach, the main aspect in nearly all identified papers is finding indicators which accurately map the extent of an ERP project. An essential question is how these factors can be ascertained. Many authors are therefore looking for parameters which can measure the actual extent of a project, including its complexity. Most researchers agree that the extent of an ERP project could not be expressed by one single parameter, but needs to be researched by way of multi-dimensional parameters.

The different assumptions of the identified models and their suggested sizing parameters will be discussed in the next chapter. Each approach has a separate sub-chapter so as to create a clean structure. First the basic principles and rationales of each approach are reported. The model is then critically assessed for its strength and limitations, and its suitability for ERP cost estimation is discussed.

Table 3.8: Identified ERP cost prediction approaches

	COCOMO II, COCOTS	Case-Based Reasoning (CBR)	Social Choice Theory	Data Envelopment Analysis (DEA)	Activity-based ERP Effort Estimation	Construct of Organisational Integration	Regression and Correlation Analysis
(Myrtveit & Stensrud, 1999b)		X					X
(Stensrud, 2001)	X	X					X
(Francalanci, 2001)							X
(Barki & Pinsonneault, 2002)						X	
(Widmer, 2004)							X
(Kwon & Shin, 2005)		X					
(Barki & Pinsonneault, 2005)						X	
(Koch, 2007)				X			
(Equey et al., 2008)							X
Janssens et. al., 2008					X		
(Daneva, Wettflower, & de Boer, 2008)	X		X				
(Daneva, 2008b)	X						
(Daneva, 2008c)	X						
(Koch & Mitlöhner, 2010)		X	X	X			X
Frequency	**4**	**4**	**2**	**2**	**1**	**2**	**6**

3.3.1 TRANSFERABILITY OF ESTIMATION MODELS FOR CUSTOM-BUILT SOFTWARE (COCOMO II & COCOTS)

The systematic literature review found two authors, namely Stensrud and Daneva, dealing with the transferability of COCOMO II and COCOTS to ERP cost estimation.

Parametric cost estimation models have their origins in individual or custom-built software and are generally defined as "a technique that develops cost estimates based upon the examination and validation of the relationships which exist between a project's technical, programmatic, and cost characteristics as well as the resources consumed […]" (Eck, Brundick, Fettig, Dechoretz, & Ugljesa, 2009).

As mentioned earlier, contrary to ERP packages, custom-built software is designed to fit the specific needs of one organisation and is developed exclusively to achieve this. In order to predict the costs of such developments, organisations need to determine the extent of their required programming effort.

In order to estimate the effort, which is mostly done in person-months, the technique establishes relationships between the size of the software project and the ability of the programmers. The ability of the programmers is measured with the unit "productivity".
The crucial point is the determination of the project size.
Scaling the software size can be conducted by way of different approaches. The most common approaches are measurement of the number of source lines of codes (SLOC), the number of function points or the number of object points. Therefore, the generic prediction estimation of custom-built software is calculated using the following scheme (Stensrud, 2001):

$$E = A \times X^{B}$$

E: Is the dependent variable effort, often measured in person-months.

A: Is a constant, representing the average productivity of the involved programmers, often determined through former software projects.

X: Is the independent variable, often expressed by source lines of code (SLOC), function points or object points (Eck, Brundick, Fettig, Dechoretz, & Ugljesa, 2009; Stensrud, 2001)

B: Is the exponential factor which represents economies or diseconomies of scale.

The most prominent model of effort prediction for custom-built software is COCOMO II. COCOMO II computes the person-months of the software development by considering the four aspects: “project size”, “productivity of the programmer”, “effort multipliers” and “scale factors”.

The size of a project is estimated by counting the required object points. Since those points differ in their complexity, the object points are converted to the required source lines of code (SLOC) of the new software. The calculation base is shown below:

$$PM = A \times Size^{E} \times \prod_{i=1}^{n} EM_{i}$$

PM: Person-Month

A: Productivity Constant

Size: Kilo SLOC (KSLOC)

E: Sum of Scale Factors

EM: Effort Multipliers

Analyst capability, Personnel continuity, Programmer capability, Application experience, Database size, Documentation, Platform volatility, Product complexity, Program language and tool experience, Required implementation schedule, REUSE, Use of software tools

At the time, commercial pre-built software packages became an important component in the design of new software systems, and COCOMO II model was stretched to its limits. That is why Boehm expands his model by including commercial-off-the-shelf (COTS) elements. In cooperation with Abts and Clark, Boehm developed the so-called Constructive Commercial-Off-The-Shelf (COCOTS) model which, in contrast to COCOMO II, aims “to predict the […] costs of using COTS components by capturing the more significant COTS risks in its modelling parameters” (C. Abts, Boehm, & Clark, 2000, p. 1)

To calculate the total effort of COTS software integration, COCOTS proposes to sum up the efforts of the four variables “assessment effort”, “tailoring effort”, “glue code effort” and “system volatility effort” as presented below:

Total Cots Integration Effort = Assessment Effort + Tailoring Effort + Glue Code Effort + System Volatility Effort

The calculations of each variable proposed by the authors are presented in Appendix II. The approaches to transfer these parametric models for custom-built software to ERP are reported in the next section.

3.3.1.1 PRESENTATION OF MODEL TRANSFERABILITY

As mentioned above, the systematic literature review identified Stensrud and Daneva as the only authors dealing with the transferability of parametric effort prediction models to ERP systems.

Stensrud (2001) theoretically compares the suitability of different existing effort prediction systems used in custom-built software for ERP cost prediction. Two of the discussed models are COCOMO II and COCOTS. Stensrud argues that both models have predefined input variables which are not appropriate for ERP projects.
According to him, COCOMO II relies on unsuitable effort multipliers which are hard to transfer. In his opinion the included multipliers “analyst capability”, “personnel continuity”, “programmer capability”, “application experience”, “database size”, “documentation”, “platform volatility”, “product complexity”, “program language” and “tool experience”, “required implementation schedule”, “REUSE”, “use of software tools” rather concern software development characteristics which differ fundamentally from ERP implementations. Consequently, Stensrud evaluated COCOMO II as not suitable for ERP cost prediction.

The same argument is applied to his evaluation of COCOTS. According to him, the predefined input variables of COCOTS – “experience with product”, “personnel capability”, "experience with COTS integration process", "personnel continuity", "product maturity", "supplier extension willingness", "product interface complexity", "supplier product support", "supplier provided training & documentation", "constraints on application system/ subsystem reliability", "application interface complexity", "constraints on COTS technical performance" and "application system portability" are mainly related to programming and not to ERP. In consequence, Stensrud found this model not suitable for being transferred to the prediction of ERP costs.

Contrary to Stensrud, Daneva (2007, 2008a, 2008b, 2008c), Daneva & Wieringa (2008), and Daneva, Wettflower, & de Boer, (2008) saw a good starting point in COCOMO II. Her study focuses on the issue of calculating the uncertainties of parametric values of the cost drivers. She developed a model which combines the traditional COCOMO II concept with portfolio management, aiming to determine "whether the use of portfolio management increases the chance of success and, if so, to what extent" (Daneva, 2008, p. 148). She conducted a case study analysis and found that the application of portfolio management increases the probability of success. However, since this research approach has a different research focus, it will not be considered in this thesis.

3.3.1.2 CRITICAL APPRAISAL

COCOMO II was designed in the context of custom-built software development and should predict the required person-months of new software. This custom-built software is mainly about programming, which is reflected in its calculation formula.
The presented formula focuses mainly on the average productivity of the involved programmers (constant A) and the project size measured in source lines of codes (SLOC). This is possibly the most significant feature that makes it difficult for COCOMO II to be transferred to ERP cost predictions.

Moreover, the author of this thesis follows Stensrud's opinion that the included effort multipliers are not suitable for effort prediction of ERP projects.
Of course some of the included multipliers are meaningful also for ERP prediction. For example, the included "soft" multipliers "analyst capability, "personnel continuity, "pro-

grammer capability" and "application experience" are likely to be an issue in ERP projects. The technical effort multipliers "product complexity" and "implementation schedule" are also expected to have an impact on ERP effort.
But the remaining technical multipliers "database size", "documentation", "platform volatility", "program language and tool experience", "REUSE" and "use of software tools" seem to be more significant for custom-built software and are hardly transferable to ERP prediction.

However, all of these multiple variables do not map the salient feature of ERP projects. Since ERP consists of packaged and standardised software, the major challenge is to integrate the so pre-defined processes into the workflow of the specific organisation (Bingi, et al., 1999; Brehm, et al., 2001; Holland, et al., 1999). There are relatively few comparisons of the programming efforts and costs in ERP projects and custom-built software.

Stensrud (2001) and Francalanci (2001) also emphasise this difference to traditional software developments. Stensrud (2001, p. 414) argues that ERP projects need to be estimated by using a multi-dimensional 'project size' instead of predicting them with a "single size measure such as function points (FP) or source lines of code (SLOC) as the main predictor variable". In contrast, measuring the size of ERP projects should include the "count of the following: users, sites, business units or legal entities, software interfaces, EDI interfaces, data conversion software and data conversion, custom-developed reports, modified screens, and ERP modules" (Stensrud, 2001, p. 414).

This is a very important aspect. The pre-defined set of input variables allows considering just a segment of the whole ERP project and giving those pre-defined variables a certain effect, small as it may be in reality.
Consequently, Stensrud (2001) identified the transferability of COCOMO II to ERP cost estimation as not suitable. The author of this thesis follows this opinion absolutely because of the given arguments.

In contrast to COCOMO II, COCOTS was developed to estimate the integration of COTS elements into the software system. Although COCOTS regards the aspects assessment, tailoring, glue code and volatility, which are indeed very important factors of ERP projects, COCOTS also considers just a pre-defined set of input variables. According to Stensrud (2001, p. 422), other important aspects like "users, sites, business units, interfaces, data conversion, custom reports, modified screens and ERP modules" are completely ignored. In conclusion, COCOTS is also not suitable for providing reliable results in ERP cost estimation, but helpful in providing ideas for new concepts (Stensrud, 2001).

3.3.2 CASE-BASED REASONING

Basically, case-based reasoning is a problem-solving paradigm that originated in the discipline of artificial intelligence. It has grown from a rather small and specified research field in the early 1980s to popularity in several research areas (Aamodt & Plaza, 1994). The central idea behind case-based reasoning is to discover similarities between new problems and previous problems in order to predict the result by using analogy (Patnaik, Malhotra, & Sahoo, 2004). In other words, this approach assumes the existence of parallelism between previous and future problems and uses old experience in order to understand and solve new problems (Kolodner, 1992). For this reason, case-based reasoning is often also called analogy-based reasoning.

Case-based reasoning became a point of interest for software development cost estimation in the 1990s. Previously, major attention was given to parametric effort models such as COCOMO II (Patnaik, et al., 2004; Shepperd & Schofield, 1997; Shepperd, Schofield, & Kitchenham, 1996; Walkerden & Jeffery, 1999). Since these models showed limitations, case-based reasoning established an alternative approach to these algorithmic effort estimation models.

The idea behind cost estimation by analogy is storing completed projects in a database. When a prediction of a new project is required, the database will be searched for the most similar project having the most similar project characteristics.
This is exemplarily shown in table 3.9 for a fictive project in which the costs of a project X should be estimated. To highlight the principle of case-based reasoning, the important aspects of the actual project are compared to the aspects of former ones.

Table 3.9: Previous projects A, B, C and the new project X

Project	Aspect 1	Aspect 2	Aspect 3	Effort
X	15	30	10	?
A	15	30	10	500
B	14	25	11	400
C	14	27	12	420

This example shows the existence of one similar project (A). Therefore the effort can be predicted by transferring the effort of project A to project X.

Since there are many more aspects than the three exemplary ones, the probability of finding identical former projects is low. There are different ways of building up case-based reasoning systems. The most prominent ones are the computer-based programmes ESTOR, developed by Mukhopadhyay et al. (1992), ANGEL[10] by Shepperd et. al. (1996), and ACE[11] by Walkerden & Jefferys (1999) .

All three methods use the Euclidean distance metric to measure the similarity to previous projects. The Euclidean distance is ascribed to the Pythagorean Theorem and represents the length of the line segment connecting the two points P (e.g. project B) and Q (e.g. project X) of an n-dimensional space.
If not all project characteristics should be weighted with equal importance, a weight for each project characteristic can be applied. The weighted Euclidean distance then has the following formula (Keung, 2009):

$$d(P,Q) = \sqrt{\sum_{i=1}^{n} w_i (q_i - p_i)^2}$$

The case (project) within a case base can be ranked according to the least Euclidean distances. The case with the lowest Euclidean distance is the most similar project.

---

[10] **AN**alo**G**y softwar**E** too**L** (ANGEL)

[11] **A**nalogical and Algorithmic **C**ost **E**stimator (ACE)

The differences between ESTOR, ANGEL and ACE can be found in the case adoption process and the number of source analogues (Keung, 2009). In contrast to ANGEL and ACE, the ESTOR tool uses pre-defined input project characteristics, which are the "number of function points" and the effort multipliers identified in the COCOMO model.
ANGEL and ACE do not need pre-defined input variables (Keung, 2009; Walkerden & Jeffery, 1999). The important aspects can be self-defined, in consequence.

Anyway, the central question regarding the transferability to ERP cost estimation is the selection of suitable aspects or input variables which need to be compared for similarity. To date, case-based reasoning has been applied in four different studies within the ERP cost estimation context. Three of them are related to ANGEL; the fourth one created a new model based on the rationale of case-based reasoning. This is why this chapter has two sub-sections, as follows.

3.3.2.1 MODEL PRESENTATION ANGEL

The application of analogy-based ERP effort estimation approaches is rarely researched. The systematic literature review found three authors who assessed the issue applying the ANGEL approach.

The first study was released by Myrtveit & Stensrud (1999b). They empirically compared the results attained by using the analogy-based tool ANGEL to their findings applying regression models.
In both approaches they regarded the following ten aspects: "number of users", "number of sites", "number of plants", "number of companies", "number of interfaces", "number of EDIs", "number of conversions", "number of modifications", "number of reports", and the "number of modules".
Myrtveit & Stensrud tested the consideration of theses aspects on 48 SAP R/3 projects and compared their findings by using the quality indications MMRE and MdMRE metric.

$$MRE = \left| \frac{(Actual\ Effort - Estimated\ Effort)}{Actual\ Effort} \right| \times 100$$

$$MMRE = \frac{1}{n} \sum_{i=0}^{n} MRE_i$$

They concluded that the analogy-based tool ANGEL provides poorer results than regression analysis. The results of their comparative study are presented in Table 3.10.

The second empirical study by Koch & Mitlöhner (2010) presented a contradictory result. In their study, the authors compared the suitability of ANGEL to social choice, data envelopment analysis and regression. Due to the comparison of MMRE, they identified ANGEL to be an essentially better approach than regression analysis. This is presented in the table below:

Table 3.10: Findings case-based reasoning compared in MMRE

Method	Author	MMRE in %
Multiple Regression Analysis	(Koch & Mitlöhner, 2010)	1159
	(Myrtveit & Stensrud, 1999)	127
Case-based Reasoning (ANGEL)	(Koch & Mitlöhner, 2010)	48
	(Myrtveit & Stensrud, 1999)	154

In contrast to Myrtveit & Stensrud, Koch & Mitlöhner chose just five different aspects for building an analogy from former projects to the new one. They selected the “number of users”, “number of locations”, “number of interfaces”, “number of modifications in LOC” and “number of modules”. Table 3.11 highlights this difference.

Table 3.11: Aspects concerned in case-based reasoning approaches

Project Characteristic	(Myrtveit & Stensrud, 1999)	(Koch & Mitlöhner, 2010)
No. of users	x	x
No. of sites /locations	x	x
No. of plants	x	
No. of companies	x	
No. of interfaces	x	x
No. of EDIs	x	
No. of conversions	x	
No. of modifications	x	x
No. of reports	x	
No. of modules	x	x

Koch & Mitlöhner tested the mentioned aspects on the basis of a dataset including 39 completed ERP implementation projects. Since they compared their results with the three quality criteria MMDR, MdMRE and pred (0.2), which attained different results, the data interpretation depends on the regarded quality criterion.

The authors conclude that ANGEL is able to generate mediocre results in comparison to the other approaches data envelopment analysis (DEA), social-choice theory and regression. However, with regard to the MMRE of the results, the analogy-based estimation ANGEL provides the best results in its comparative analysis.

The third identified study published by Stensrud (2001) is already mentioned in the COCOMO II section. He compared and assessed different existing effort estimation approaches theoretically, among them ANGEL, ESTOR and ACE.

He concluded that regression analysis is the most suitable approach for making reliable and accurate predictions of ERP effort. He explained that with the non-parametric nature of ANGEL.

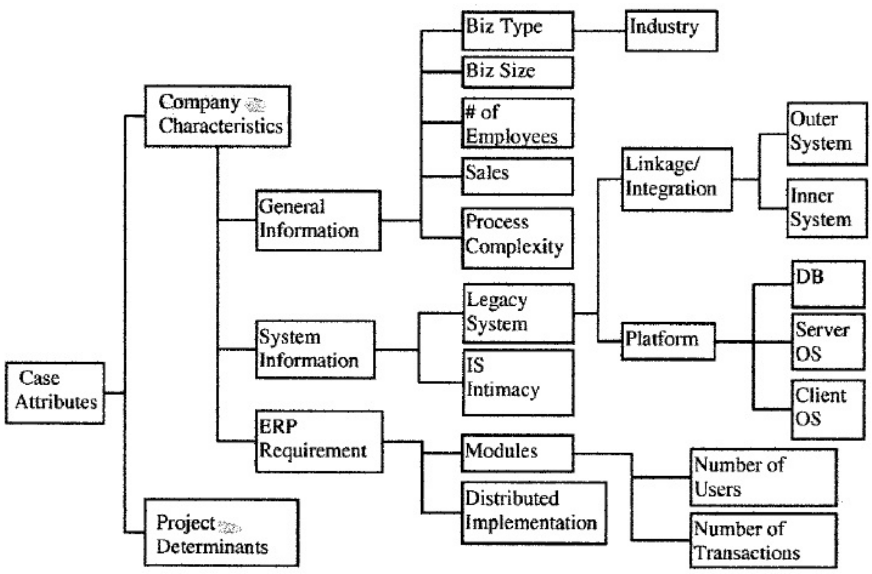
In his opinion, those models are not able to provide a relationship between the project size and the effort. According to him, this can only be accurately achieved with regression analysis. Nevertheless, Stensrud concluded that ANGEL and regression analysis should be used complementarily. According to him (p. 422 et seq.), “Regression analysis assists the user in understanding the historical data by synthesising the data, i.e. by data reduction, whereas ANGEL adds value by drilling down in the data".

3.3.2.2 MODEL PRESENTATION PROJECT PRE-PLANNING SUPPORT SYSTEM (PPSS)

This sub-chapter introduces the fourth paper dealing with case-based reasoning.

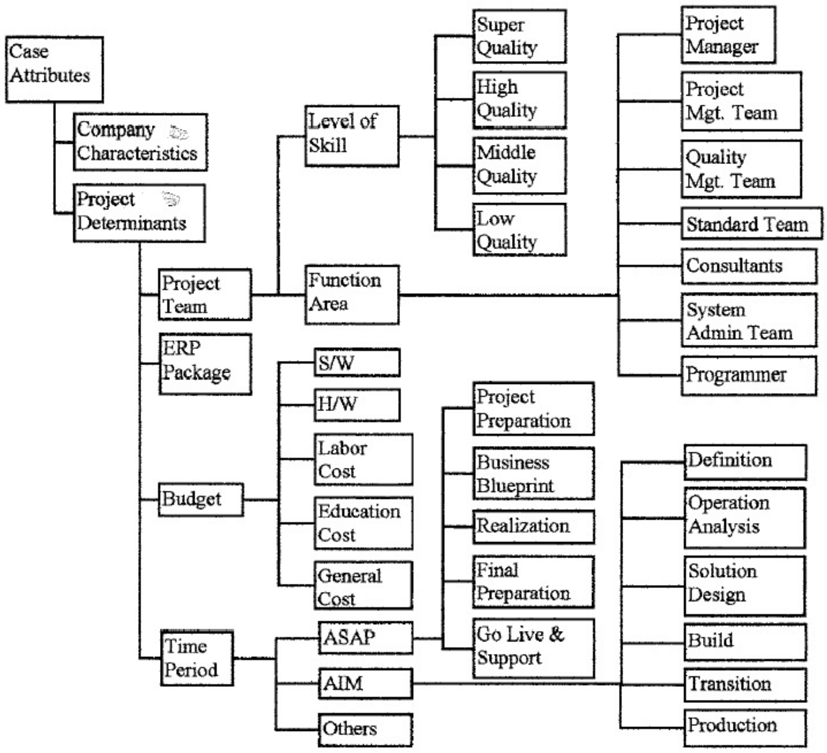
In 2005, Kwon & Shin developed a two-step concept for predicting the project costs of ERP systems. Their computer program, the so-called 'Project Pre-planning Support System' (PPSS), finds the most similar case according to the CBR rules in a first step and allows a rule-based reasoning for the adoption in the actual projects in a second step. With this approach, Kwon & Shin aim to enable ERP managers to carry out more precise cost planning for their projects (Kwon & Shin, 2005).
Kwon & Shin designed a selection scheme in order to initially find the most similar case compared to current case. This scheme is composed of the two main categories 'company characteristics' and 'project determinants'.
These aspects and their respective sub-categories are presented in the following two figures, Figure 3.1 and Figure 3.2:

Figure 3.1: Company characteristics attributes



Source: (Kwon & Shin, 2005)

Figure 3.2: ERP project determinant attributes



Source: (Kwon & Shin, 2005)

Figure 3.1 shows the categories of the main company characteristics as “general information”, “system information”, and “ERP requirement”. Their attributes are then elaborated into more specific attributes, resulting in a total of 14 company characteristics (industry, business size, no. of employees, sales, process complexity, information system intimacy, distributed implementation, order system, inner system, database, server operating system, client operating system, number of users, and number of transactions).

Figure 3.2 presents the project determinants consisting of four main categories, which are “project team”, “ERP package”, “budget”, and “time period”. These four main project determinants are then also elaborated into more specific attributes, resulting in a total of 29 project determinants (ERP package, software budget, hardware budget, labour cost, education cost, general cost, others, super quality, high quality, medium quality, low quality, project preparation, business blueprint, realization, final preparation, go live & support, project manager, project management team, quality management team, standard

team, consultants, system administration team, programmer, definition, operation analysis, solution design, build, transition, and production)
Altogether the PPSS concept provides 43 attributes for the sizing and classification of ERP cases.

The similar size of cases is calculated by applying the nearest neighbour function using the following equation:

$$Similarity(T,S) = \sum_{i=1}^{n} f(T_i, S_i) \times w_i$$

T: Target case
S: Source case
n: Number of attributes in each case
i: Individual attribute from 1 to n
f: Similarity function for attribute i in cases T and S
w: Importance weighting of attribute i

The variable w permits to weight aspects in order to "reflect the comparative importance" of aspects (Kwon & Shin, 2005, p. 164).
After PPSS identifies the most similar case, rule-based reasoning allows an adjustment of the findings to the current case. Therefore, the rule bases designed by Kwon & Shin are applied. According to the authors, rule-based reasoning establishes rules which follow the material conditional. It logically connects statements to an implicated relationship which is usually expressed as "p→q" and generally termed as an "if p than q" relationship.

Kwon & Shin designed their rule base by conducting interviews with experienced ERP managers and deriving relationships from their findings. They gave an example for one rule which states "If ERP is implemented at multiple places and to be integrated, cost for consulting manpower usually rises by 10 – 20 %". Although the author claims that there "exist some causal relationships between aspects" (Kwon & Shin, 2005, p. 164), the reader finally remains unclear about the applied relationships between aspects and the so developed rules. Since Kwon & Shin do not mention a quality criterion, like MMRE, an assessment of their program his hardly possible.

However, Kwon & Shin close their research by declaring PPSS a prototype containing eight cases to date. In order to provide a real benefit to ERP managers, a lot more cases are needed.

3.3.2.3 CRITICAL APPRAISAL

The two previous sub-sections presented the case-based models applied in ERP effort estimation. Their suitability for predicting ERP costs will be discussed in the following.

The key aspect in case-based reasoning is to find similarities between the project to be estimated and projects completed in the past. The main question is what the correct aspects for a similarity comparison are. Since ERP projects vary strongly from one organisation to another, finding the key aspects for comparing similarities is essential to the accuracy of the whole model.
Focusing on the wrong similar aspects can be very misleading. In consequence, ERP project managers must have a basic notion of the essential project characteristics and their loading.

The case-based reasoning approaches discussed in ERP cost prediction make different suggestions about this key question.
Myrtveit & Stensrud (1999b) propose these 10 aspects: “number of users”, “number of sites”, “number of plants”, “number of companies”, “number of interfaces”, “number of EDIs”, “number of conversions”, “number of modifications”, “number of reports” and “number of modules”. Koch & Mitlöhner (2010), in contrast, consider just these five different aspects: “number of users”, “number of locations”, “number of interfaces”, “number of modifications” and “number of modules”. And Kwon & Shin (2005) presented a selection scheme composed of the two main categories ‘company characteristics’ and ‘project determinants’ which include numerous aspects.

However, Myrtveit & Stensrud (1999b) achieved better results applying regression analysis than with the analogy-based tool ANGEL. Koch & Mitlöhner (2010) yielded completely opposite results in their comparative study. Unfortunately, Kwon & Shin (2005) did not present any quality criterion that enables an assessment of their programme.
Maybe the "right" aspects are not found in the end and project managers have no guarantee of considering all relevant aspects when predicting their ERP cost.

In 2001, Stensrud critically assessed the suitability of approaches for ERP effort estimation and identified the main weakness of ANGEL in the non-parametric nature of case-based reasoning. Since ANGEL is not able to establish a relationship between the size of a project and its effort, he saw the value of this model rather in exploratory data analyses than in effort prediction (Stensrud, 2001).
According to Stensrud, the suitability of transferability of the other programs ACE and ESTOR can be evaluated similarly to the suitability of ANGEL.
Like ANGEL, ACE has no pre-defined input variables. This means the problem of choosing the right aspects for measuring similarity remains. ESTOR uses a pre-defined set of project characteristics which focus on software development aspects but do not meet the characteristics of ERP projects (Stensrud, 2001). Stensrud (2001) concluded that ESTOR is consequently not suitable to be transferred to ERP cost estimation.

The author of this thesis agrees with Stensrud's assumptions and these limitations. Nevertheless, case-based reasoning yields good results, apparently. It seems to be a promising approach that provides helpful indications.

After discussing the case-based reasoning model ANGEL, the suitability of the PPSS model designed by Kwon & Shin (2005) should be examined.
The idea behind this model is to find the best matching case for a new project in the first step, and to adjust the findings to the actual circumstances in a second step. Adjusting is done by applying Kwon & Shin's rule bases which follow the material conditional.

Although adjusting the most similar project to the current project is assuredly a good concept, the valuation of the designed rule base is rather difficult. Since Kwon & Shin

(2005) leave their readers uncertain about their applied rules and lack a quality criterion, the suitability of the applied rule-base reasoning cannot be assessed.

A major weakness of the approach is the quantity of completed cases deposited in the database. Kwon & Shin (2005) admit that a database consisting of eight cases can hardly be more than a prototype. The issue with the database is one thing that needs to be generally discussed. Since there are no publicly accessible databases which store completed ERP cases, the prediction of ERP costs according to similarities must be assessed as rather limited.

Myrtveit & Stensrud (1999b) refer to a database containing 48 SAP R/3 projects, the results of Koch & Mitlöhner (2010) are based on 39 completed ERP implementation projects, and Kwon & Shin (2005) reference 8 cases.

It must be questioned if the quantity of cases is enough to provide good results and, in the case of Myrtveit & Stensrud, if the SAP projects can be transferred to other ERP vendors. Building a more extensive database seems to be a good idea anyway.

3.3.3 SOCIAL CHOICE THEORY

Social choice theory generally focuses on systems or institutions aggregating preferences to a collective choice (Kelly, 2013). Democratic elections are probably the most famous research issue. Social choice assumes that the "resolution of […] a group choice or collective action should be based on the desires or preferences of the individuals in the society, group or collective." (Fishburn, 1973, p. 3).

To map the choice set and the preferences of individuals over a set of alternatives in abstract terms, the discipline developed social choice function.

Basically, N is the quantity of individuals within the society, group or collective who have to select from a choice set X of social alternatives (Endriss, 2011).

An example highlights how individual preferences could be aggregated to a collective preference: In this example, the two voters $N=\{1,2\}$ rank their preferences over a choice set of three alternatives $X=\{A,B,C\}$.

Table 3.12: Example for aggregation of individual preferences

Preference	**Voter 1**	**Voter 2**
1st Preference	B	A
2nd Preference	C	B
3rd Preference	A	C

The preference of voter 1 can be mapped with the following function: $B \succ C \succ A$.
This indicates that B is preferred to C, whereas C is at once preferred to A.
The choices of voter 2 can be expressed as $A \succ B \succ C$.
This means that A is preferred to B, whereas B is at once preferred to C.

It is also possible that two or more alternatives are equally preferred. In case the alternatives A and B are equally preferred, but both preferred to C, this would be indicated as follows: $(A \sim B) \succ C$

However, there are different voting rules for aggregating such preferences within the social choice theory. The most famous are e.g. 'simple majority', 'Borda count', 'Condorcet social choice function' etc. (Fishburn, 1977; Klamler, 2005; Risse, 2005; Young, 1995).

The 'simple majority' voting rule declares the winner of an election by the highest number of 1st preferences only.
The 'Borda count' voting rule requires a ranking list of candidates from every voter and weighs the rank with a defined score. In a three-candidate election, for example, the first preference will be scored with 3 points, the second one with only 2, and the third preference with 1 point. Finally, the number of points will be summed up and the total yields a collective ranking list based on the total score of every candidate.
The 'Condorcet method' uses a pairwise comparison between every candidate. Similar to the 'Borda count' every voter must generate a ranking list according to her/his preferences. The 'Condorcet method' then compares pairwise how often a candidate is ranked above her/his opponent (Sen, 1986).

The dependence between electoral results and the aggregating rules is very essential. This should be highlighted by the following example. Table 3.13 shows the possible combinations of a choice set including three alternatives $X = \{A, B, C\}$ and the stated preferences of 21 voters $N = 21$.

Table 3.13: Example of preferences of choice set including three alternatives

	Combi 1	**Combi 2**	**Combi 3**	**Combi 4**	**Combi 5**	**Combi 6**
1st Preference	A	A	B	B	C	C
2nd Preference	B	C	A	C	A	B
3rd Preference	C	B	C	A	B	A
# Of Stated Preferences	**6**	**0**	**5**	**2**	**5**	**3**

Adapted from Wahltheorie (2011): http://www.math.uni-potsdam.de/prof/o_didaktik/aa/zz01/sozialwahltheorie.pdf

Applying the 'simple majority' voting rule, candidate C would be the winner with eight stated first preferences. Candidate B with seven first preferences would have the second place and candidate A with six first preferences the last place.
Results will differ strongly when the preferences are aggregated using the 'Borda count' method[12]. In this case, candidate B is the winner with a total of 44 points. Candidate A will be second with 43 total points before candidate C with 39 points.

When using the 'Condorcet method' a third result is yielded in which candidate A is the winner over candidates A and B, whereas candidate B is before candidate C.

$A \succ B$: 6+0+5 = 11 $\quad$ $B \succ A$: 5+2+3 = 10 $\quad$ → $A \succ B$

$A \succ C$: 6+0+5 = 11 $\quad$ $C \succ A$: 5+2+3 = 10 $\quad$ → $A \succ C$

$B \succ C$: 6+5+2 = 13 $\quad$ $C \succ B$: 0+5+3 = 8 $\quad$ → $B \succ C$

→ $A \succ B \succ C$

[12] In this example, the following scoring rule was applied: Candidates' votes equipped with a 1st preference are scored with three points. 2nd preferences are scored with two points, and 3rd preferences are scored with only one point.

This example shows that the aggregation rules are able to yield totally different results. They play an important role in preference accumulation and are analysed carefully by social choice researchers.

Despite that, social choice theory stands for bringing preferences into a sorted order. It does not involve a quantification of preferences and their distances to each other, which leads to a “relative assessment”. A qualitative or absolute valuation of the alternatives is not provided. Only a classification in relation to the other alternatives is conducted, and distances between preferences are disregarded.

Due to this characteristic, social choice theory is basically open to bringing diverse characteristics, even outside of the subject area of elections or votes, into a ranking order. Because a value-based assessment is not required and only a comparative assessment with “$\succ$”, “$\prec$”, or “$\sim$” is necessary, this technique could be used as a simplified method for measuring similarity when assuming that a closer rank equals a higher similarity.
In contrast to case-based reasoning, social choice does not require absolute values. Instead, the cases only have to be brought into a ranking order.
The application of social choice theory to the topic of ERP estimation will be presented in the following sub-section.

3.3.3.1 APPROACH PRESENTATION

The systematic literature review identified just one application of social choice theory in the ERP cost estimation literature. This is the comparative study of Koch & Mitlöhner (2010) mentioned in the previous section. They compared the suitability of social-choice theory with other approaches in an empirical study.

Koch & Mitlöhner (2010) adjusted the social choice theory to ERP cost estimation by designing a ranking and weighting system for ERP characteristics. Koch & Mitlöhner (2010) suggest to replace the “individual preferences” with ERP project characteristics and “alternatives” with a database of completed projects from the past.

According to Koch & Mitlöhner, each numeric project characteristic should be equipped with a ranking that considers all projects from the database. The rankings imply the project size and complexity of the projects according to one characteristic. The aggregated ranking of all attributes will then “represent an overall picture of the projects ‘complexity’, and thus a ranking of the efforts necessary to implement them.” (Koch & Mitlöhner, 2010, p. 269).
Therefore, it is necessary that the implementation efforts of the projects within the database are known.

In order to estimate the effort of a new ERP project, the project must be ranked in reference to the database projects. This needs to be done for every project characteristic. It is important to understand that the absolute value parameters of the characteristics of the new project do not have to be known. Every characteristic is just ranked between the other projects.

Since each considered project characteristic might have a different degree of importance, Koch & Mitlöhner developed a loading system in which each characteristic is weighted between the numbers 0 and 99. This weighting was to be considered in the aggregating procedure.

Aggregating the ranking of the characteristics, the new project to predict will allot a place between other ones, of which the costs are known. An indication of the effort of the new project is provided based on the known effort values of the neighbours in that ranking list (Koch & Mitlöhner, 2010).

Koch & Mitlöhner (2010) use the following five project characteristics in their study:
- Number of users
- Number of locations
- Number of interfaces
- Number of modifications in LOC
- Number of modules

Since aggregating rules can have a deep impact on the results, Koch & Mitlöhner tested different aggregating rules in their study, yet they found only small variations in the results.

According to the authors, 'social choice model' produces quite good results. Regarding the quality criterion pred (0.2), they identified that their social choice approach outperforms all other estimation models examined in their comparative study, like data envelopment analysis (DEA), analogy-based estimation, classification, and regression trees and multiple regression. Another picture is presented when considering MMRE as a quality indicator. With an MMRE of 543% (Copeland) and 958% (Borda), the results appear to be quite poor.

3.3.3.2 CRITICAL APPRAISAL

The basic idea of this approach is to evaluate and estimate the effort of a new project in comparison to completed ones. This part of the approach is quite similar to the analogy based models. For this reason, social choice offers the same points of advantages and of criticism as the analogy-based approaches presented in the previous section.
In contrast to case-based reasoning, social choice does not require absolute value parameters of the new project. Instead, the cases only have to be brought into a ranking order and arranged between two completed database projects.
Although this seems to be an advantage at first view, the realisation of this ranking system in practice seems to pose some difficulties. It appears hard to image how a ranking should be done without having an idea of the value characteristics.

Koch & Mitlöhner suggest using the following five project characteristics for an evaluation:
- Number of users
- Number of locations
- Number of interfaces
- Number of modifications in LOC
- Number of modules

Although the "number of users", "number of locations" and "number of interfaces" can easily be ranked, an assessment of the "number of modifications" and "number of modules" appears to be rather difficult in advance.
Since the features within a module vary strongly from one vendor to another, as mentioned in chapter 2.2.3, a reliable ranking of this project attribute seems to be only possible after choosing the supplier. The "number of modifications" seems to be an unpredictable attribute which is very hard to predict in advance. A misranking could have fatal impacts on the effort.

However, another question is if Koch & Mitlöhner selected the "right" project attributes for providing reliable results. When searching for similar projects, one would likely look for projects which used the same vendor with the same modules. This is not done in this approach.

Another issue is that only technical project characteristics are regarded. A consideration of further influences, such as organisational or situation-related quantities, is not given any attention in this approach. Factors like the number of consultants are not included in the model. Since there is no clarity about the actual relevant factors for estimating ERP effort, some uncertainty about the "right" project attributes remains.
In principle, however, this model is open to more than the suggested characteristics and could be extended in future research.

The social choice approach basically provides a good reference point. But, like case-based reasoning, the approach cannot explain the relationship between project size and project effort. Since it has only been applied in one study to date, the approach needs to be tested in future research to make an appropriate evaluation about the strengths and weaknesses of this approach. Unfortunately, the high MMRE is an indicator that the project characteristics regarded so far are not able to provide reliable results.

3.3.4 DATA ENVELOPMENT ANALYSIS

The systematic literature review identified two papers dealing with data envelopment analysis (DEA) for ERP cost estimation.

The data envelopment analysis (DEA) model was developed by Charnes, Cooper & Rhodes in 1978. It is a non-parametric approach that compares the relative efficiency of so-called decision-making units (DMUs), such as organisations, in order to investigate their improvement potential (Koch, 2007; Myrtveit & Stensrud, 1999a; Sherman & Zhu, 2006). To analyse their relative efficiency, DEA investigates the transformation of multiple inputs into multiple outputs, which is defined as the ratio of outputs to inputs. This is represented by the following formula:

Efficiency = Output / Input

The highest possible ratio of outputs to inputs is called efficiency frontier.

Efficiency can generally be increased by raising the output (output efficiency) or decreasing the input (input efficiency). Output efficiency focuses on the increase of efficiency by enhancing the outputs and keeping the inputs constant. Input efficiency is the increase of efficiency by reducing the inputs and keeping the outputs constant (Koch, 2007; Sherman & Zhu, 2006). In consequence, there are generally two alternative ways to determine the efficiency: the input reducing or output increasing efficiency (Stensrud 1999b).

However, DEA usually looks at more than one input and one output variable; it generally allows for considering several input factors and several different output factors. Since the outputs and inputs cannot be added up directly (Camanho, 2007), it is necessary to weight the inputs and outputs by their relative values (Sherman & Zhu, 2006). DEA does that with the following formula:

$$E_k = \frac{\sum_{r=1}^{s} u_r \times y_r}{\sum_{i=1}^{m} v_i \times x_i}$$

E_k Efficiency of DMU k

u_r Coefficient, weight assigned to output r

y_r Amount of output r

v_i Weight assigned to input i

x_i Amount of input i

DEA has the ability to determine these weights automatically by using computed linear programming techniques that allow the determination of the highest possible efficiency ratio, since the programme maximises the "set of coefficients [u's and v's] that will give the highest possible efficiency ratio of outputs to inputs" (Sherman & Zhu, 2006, p. 63).

The result can vary between 0 and 1. An efficiency rating of 1 is the highest possible efficient rate, whereas a result of 0 is the most inefficient rate (Sherman & Zhu, 2006).

3.3.4.1 APPROACH PRESENTATION

The systematic literature review identified two papers considering the suitability of DEA for ERP cost estimation. This is the work of Koch (2007) and the comparative study of Koch & Mitlöhner (2010) mentioned several times.

Koch (2007) empirically examined the idea of applying DEA for predicting the effort in ERP projects. He used an output increasing efficiency measure.
He used "total costs" as input variable. As output variables, he suggested "duration", "number of users, "number of interfaces", "number of modules", "extent of modifications" and "number of locations" as potentially suitable.
In order to identify the "right ones", he tested his suggestions for their correlation with the input variables in his data set, which consisted of 39 Australian companies.
He excluded "extent of modifications" and "number of locations" as possible output variables because they did not show any significant correlation.

For weighting the output variables and examining the productivity rate, he used a computed DEA program which allowed him to determine an average productivity rate of 0.718. In order to examine the quality of his DEA model, he predicted the effort in total costs for the cases of his data set using three different scenarios for the efficiency to be achieved in the project: an optimistic scenario with a productivity rate of 1, a pessimistic scenario with a rate of 0.5, and a realistic scenario with the mean value of the data set. He achieved the best results with the optimistic scenario, yielding an MMRE value of 154%.

The second paper is the already mentioned comparative study of effort estimation for ERP projects. Its authors, Koch & Mitlöhner (2010), empirically compared the DEA results Koch achieved in 2007 with social choice theory, analogy-based estimation as well as classification and regression trees, and multiple regression.
For the empirical validation they used the same data set of 39 enterprises that Koch (2007) had published in his study three years earlier.
Regarding the MMRE value of 155% in their comparative study, DEA achieved mediocre results and got third place.
The best results were attained with case-based reasoning. Poorer outcomes were attained with the applied regression and social-choice approaches.

3.3.4.2 CRITICAL APPRAISAL

The comparative study of Koch & Mitlöhner (2010) identified DEA as a method providing mediocre results.
The crucial point in DEA is finding the "right" productivity rate. Koch (2007) calculates the average productivity rate primarily as a benchmark and combines his value with a positive scenario in order to get the best results.
But the identification of this rate seems to be one major limitation in this approach. Such a rate does not allow any adoption of the project and company specific issues. Since most ERP projects are unique, the application of a general efficiency level is hardly imaginable.

Another issue is the calculation of this level. How can project managers determine their productivity rate? The average value might be a first indication, but it does not provide an accurate value for implementing organisations. Unfortunately, the approach does not present sufficient details for this determination. It remains unclear which factors are important for an increase or decrease of the productivity rate. The questions which factors are depended for an increase or decrease of the productivity rate remain unclear.

A further subject for debate is the selection of the output variables. Koch (2007) suggests that "duration", "number of users", "number of interfaces", and "number of modules" are good indicators for classifying an ERP project. It is questionable whether vast and complex ERP projects can be mapped by these attributes. The issue of finding the "right" attributes is not conclusively researched.

3.3.5 ACTIVITY-BASED ERP COST ESTIMATION

3.3.5.1 APPROACH PRESENTATION

The systematic literature review also found a theoretical approach for sizing ERP projects. Even if the current status of this approach does not provide a technique for calculating the effort, the findings could be a valuable contribution to further research.

In their paper, Janssens, Kusters, & Heemstra (2008) aimed to identify the activities within an ERP project and proposed to cluster them in order to create a basis for defining the size of the project.

In order to answer the question which activities exist, they conducted a systematic literature review and logically clustered the identified activities using the Metaplan technique. They identified 208 activities and synthesised them into 21 clusters and/or sub-clusters. Table 3.14 presents the findings and the corresponding number of included activities.

Table 3.14: ERP implementation activity clusters and sub-clusters

		Group view			
Clusters	**Subclusters**	**Project**	**System**	**Organization**	**Number of unique activities**
Selection	Vendor selection		✓		4
	Product selection		✓		16
Project configuration		✓			19
Project management	Management	✓			4
	Communication to organization	✓			4
Organizational and system design	Current state analysis			✓	5
	Organizational requirements			✓	7
	Requirements ERP system		✓	✓	8
	High level Design		✓	✓	6
Configuration and installation	System configuration		✓		17
	Data conversion		✓		4
	System integration		✓		9
	ERP system testing		✓		14
Customizing			✓		7
Infrastructure			✓		14
Reorganization				✓	11
System implementation				✓	21
Training	Training Implementation Staff	✓			2
	Training users			✓	9
	Training maintenance staff			✓	2
Set up maintenance			✓		25
TOTAL					**208**

Source: Janssens et al., (2008), p. 9

In order to categorise their findings, the authors grouped the clusters into the following three groups: “Project”, “System” and “Organisation”. According to them, the ‘project group’ contains activities which are necessary in order to keep the project running, such as project management. The ‘system group’ includes activities concerning configuration and implementation of the ERP system. The ‘organisation group’ presents activities concerning the organisation, such as business process re-engineering or training.

It needs to be researched how these findings are actually used for predicting ERP project costs.

3.3.5.2 CRITICAL APPRAISAL

As mentioned earlier, this approach does not provide a technique for actually predicting ERP effort, but delivers a detailed insight into the activities of ERP implementation projects. It remains unresolved how activities can be used for the determination of ERP size.

It must be assumed that the activities have different complexities, and it can be expected that the duration of the activities will depend on the organisation itself.
Therefore, it seems to be essential to build a kind of evaluation scale and/or weighting factors for each activity within such a model. However, it will be exciting to see how these activities may be used in further model development approaches.

3.3.6 ERP EFFORT ESTIMATION USING CONSTRUCT OF ORGANISATIONAL INTEGRATION

The literature review found two papers of the authors Barki & Pinsonneault aiming to explain ERP implementation effort by using the concept of organisational integration.

Generally, integrated organisation models focus on the interrelationships of the different elements of an organisation. If these elements do not fit each other, the organisation does not perform optimally, and a need to change becomes apparent (TACSO, 2010).

3.3.6.1 PRESENTATION OF THE CONSTRUCT

Section 2.2.1 has already emphasised the integration concept as one of the key principles of ERP systems. In their papers, Barki & Pinsonneault (2002, 2003, 2005) went even one step further, assuming the creation of a higher level of business process integration to be the biggest advantage of an ERP implementation. This led them to the hypothesis that there is a relation between the improvement of organisational integration and the effort spent on ERP implementation, which means that the ERP implementation effort[13] depends especially on the change level of organisational integration.

According to them, improvement in organisational integration is a product of "business process integration" and "technological integration".

---

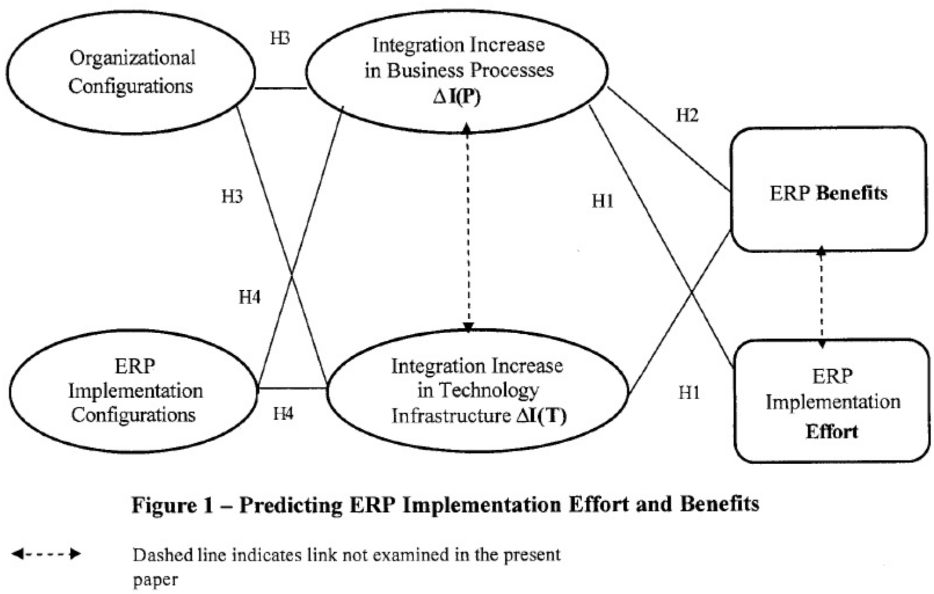
[13] Barki & Pinsonneault consider technology infrastructure costs, business process re-engineering costs, organisational change costs, and ERP system costs as ERP implementation effort.

Based on this rationale, Barki & Pinsonneault (2002, 2003, 2005) developed a hypothetical research model based on organisational integration, which should explain the ERP implementation effort by putting the implementation effort of ERP systems in relation to its benefits. Unfortunately, the model lacks empirical validation and does not present a tool for actually calculating the ERP project costs.

However, Barki & Pinsonneault (2002, 2003, 2005) suggested linking the two ERP project results "ERP benefits" and "ERP implementation effort" to the organisational integration elements, and developed hypotheses for the relations between them.

Figure 3.3 below maps their assumptions and presents which hypotheses are expected for which interrelation. The hypotheses are formulated below the figure.

Figure 3.3: Model of organisational integration for ERP effort



Source: Barki & Pinsonneault, 2002, p. 16

According to the authors, hypothesis H1 is defined as follows:

H1: The greater the change in organisational integration, $\Delta I(P)$ and $\Delta I(T)$, resulting from the implementation of an ERP system, the greater the implementation effort, E, required where:

$\Delta I(P) = I(P)_{t+1} - I(P)_t$ and

$\Delta I(T) = I(T)_{t+1} - I(T)_t$

H2: The greater the change in business process integration, $\Delta I(P)$, resulting from the implementation of an ERP system, the greater the implementation benefits, B, obtained.

The hypotheses H3 and H4 refer to the precursors of the construct of 'organisational configuration' and 'ERP implementation configuration'. H3 focuses on organisational configuration, which Barki & Pinsonneault (2002, p. 20) define as: "particular patterns in which organizational strategy, structure, environment and process variables are aligned, are a key determinant of the organizational integration of a firm prior to the implementation of an ERP system". H3 reads as follows:

H3: Organisational configurations influence the change in organisational integration, $\Delta I(P)$ and $\Delta I(T)$, resulting from the implementation of an ERP system.

'ERP implementation configuration' describes to which extent an ERP system will be implemented into an organisation. This aspect leads to the fourth hypothesis of Barki & Pinsonneault:
H4: ERP implementation configurations influence the change in organisational integration, $\Delta I(P)$ and $\Delta I(T)$, resulting from the implementation of an ERP system.

With their model and hypotheses, Barki & Pinsonneault propose a theoretical construct for assessing the effort dependent upon its benefits. These variables are solely influenced by the change of the business processes and technologies.
But, unfortunately, as mentioned earlier, they did not test their propositions empirically and do not provide a technique for determining ERP project costs in advance.

Three years later, in 2005, Barki & Pinsonneault published their second paper. In this second research, they made slight modifications to the formerly presented model. But, again, it provides just theoretical assumptions. As in their previous study, a tool to actually calculate the ERP effort is missing.

3.3.6.2 CRITICAL APPRAISAL

Barki & Pinsonneault focus on the integration aspect of ERP projects. As mentioned earlier, this issue is one of the most salient features. Their OI approach highlights the inter-organisational elements and proposes analysing the implementation effort dependent upon its benefits.

Despite including this important feature, the approach shows some weaknesses:
The most important limitation in the context of this thesis is the lack of a concrete technique or tool for estimating the effort of a project in advance. Therefore, this approach is not able to make a prediction of costs.
Furthermore, there is no validation of their theoretical assumptions. Further research is needed to determine if the suggestions can be confirmed by reality or need to be rejected. But, generally, the model raises the question how it could be applied and how integration can be measured. That seems to be a complex issue which is given no attention in this model.

The second remarkable issue is the rationale itself. Of course integration is one very beneficial aspect in ERP projects and many organisations aim to increase it by implementing ERP. But being the one and only benefit, it seems a legitimate subject for debate. There might be other objectives for implementing ERP systems, like automatically mapping the business processes, which are absolutely disregarded in this model.

Thirdly, organisation-specific issues and influencing quantities receive no consideration. Aspects, like the frequently mentioned variables “number of users”, “number of locations” or “number of modules” suggested in previous models, are not included. It is difficult to say if estimations can be made without knowing such specific properties. Empirical analysis is required to answer these questions. However, it is an overall model, and instead of seeing it as a tool, one might rather refer to it as an integration concept.

3.3.7 ERP EFFORT ESTIMATION USING CORRELATION AND REGRESSION ANALYSIS

The systematic literature review identified six papers dealing with correlation and regression analysis for cost estimation of ERP projects.

Correlation and regression analysis are very general and therefore highly flexible methods of data analysis. They are usually applied to examine and assess relationships between a number of independent variables (xi) and the dependent variable (y). Regression is often used in practical prediction problems, aiming to forecast a result based on the data collected earlier (Cohen, Cohen, West, & Aiken, 2013).

3.3.7.1 APPROACH PRESENTATION

Correlation and regression analysis is the most often applied approach for ERP effort estimation identified in the literature review. Six studies were found. They can be grouped in accordance with their research nature into "correlation analysis", "regression analysis" and "papers contributing only theoretical discussion". The classification is presented below:

Correlation Analysis

- Francalanci (2001): Predicting the implementation effort of ERP projects: empirical evidence on SAP/R3
- Equey, Kusters, Varone & Montandon (2008): Empirical study of ERP system implementation costs in Swiss SMEs

Regression Analysis

- Myrtveit & Stensrud (1999): A controlled experiment to assess the benefits of estimating with analogy and regression models
- Widmer (2004): Schätzung & Beeinflussung der Kosten von ERP-Systemen in Schweizer KMU (*Translation: Assessing & Influencing ERP System Costs in Swiss SMEs*)
- Koch & Mitlöhner (2011): Effort estimation for enterprise resource planning

Papers contributing only theoretical discussion:

- Stensrud (2000): Alternative approaches to effort prediction of ERP projects

The basic aim of most papers is to define the extent of ERP projects in order to predict the arising costs. Therefore, they attempt to find appropriate size metrics or cost drivers. The findings are presented in the sections below.

At first, the approaches using correlation analysis will be introduced, then the identified multiple regression analysis is reported, and finally, the theoretical discussion is presented.

CORRELATION ANALYSIS

The first study using correlation analysis was conducted by Francalanci in 2001. She claimed that technical parameters are not sufficient to measure the effort of ERP projects and that additional contextual factors are needed for explaining the variance for projects having the same technical size.

In her paper, Francalanci (2001) aims to measure the effort of human resources required in ERP projects. To this end, she suggests estimating the ERP effort by determining the project size and project context. According to her, the size is measured in terms of the number of modules and sub-modules, while context is expressed through the organisational size (operationalised as revenue & total number of employees) and the total number of users and per module number of users (operationalised as number of licences).

She analysed if her suggested variables correlate with the ERP effort. Her data set included 43 SAP/ R3 projects.

The result of her study presented a statistical significance of the variables. This was an important step forward since these five variables could be empirically validated.

Unfortunately, the correlation strengths are not mentioned in her paper.

The second study using correlation analysis was published by Equey et al. (2008). They aimed to investigate which cost drivers influence the ERP implementation costs.

In order to select appropriate cost drivers, they conducted a mail-based survey in Swiss companies.

In order to find appropriate cost driver candidates, the researchers conducted in-depth interviews with ERP experts. They classified these variables into three dimensions: enterprise characteristics, people and implementation. An overview of the dimensions, their grouped cost drivers and the correlation findings is presented in the table below:

Table 3.15: Correlating variables identified by Equey et al. (2008)

Dimension	Variable	Correlation strength
Enterprise Characteristic	Annual sales revenue	-0.167
	Enterprise subsidiary	-0.244
	Number of ERP users	No correlation
	Number of employees	No correlation
	Sector of activity	No correlation
People	Management involvement	0.182
	ERP consultants' level of experience	-0.241
	Employee involvement	0.171
	Ratio of external consultants to internal employees	0.172
	Employee qualifications	No correlation
	Field of expertise	No correlation
	Project manager's position	No correlation
Implementation	Number of modules	0.186
	SCM module	0.260
	Production module	0.220
	Sales /CRM module	0.274
	Inventory module	0.186
	Other types of modules	No correlation
	Organisational tool	No correlation

Although the authors report on strong relationships, the author of this thesis values the identified correlations as low or slight.
Remarkably, the authors did not find a correlation between the numbers of users. This is contradictory to the findings of Francalanci.

However, this study adds additional value to the research issue since it focuses on more cost driver candidates than Francalanci's study. Equey et al. (2008) concluded that they validated a new cost driver, which is "experience of the consultant".

Despite these findings, correlation analysis does not provide a technique for predicting future projects and can therefore not be used in ERP effort estimation at the moment.

REGRESSION ANALYSIS

The first work using regression analysis is a comparative study by Myrtveit & Stensrud (1999b). The authors compared the analogy-based approach ANGEL and regression analysis for the purpose of ERP effort estimation.

For their comparison, they initially considered ten factors for sizing the project, which are: “users”, “sites”, “plants”, “companies”, “interfaces”, “EDIs”, “conversions”, “modifications”, “reports” and “number of modules”.

For their regression model, they excluded seven of the ten variables because they were either highly correlated with other regressors, which is called multicollinearity, or not defined clearly and consistently enough.

However, the following three variables “number of users”, “number of EDIs,” and “number of conversions” remain in their analysis. Myrtveit & Stensrud (1999b) regarded them as independent variables in their comparative study. Their data set included 48 SAP/R3 projects. For assessing their results, they regarded the quality criteria MMRE, MdMRE and R^2 which yield the following outcomes:

Table 3.16: Findings regression & case-based reasoning compared in MMRE

Method	Adjusted R^2	MMRE in %
Regression	80.1	127
Analogy-based estimates	-	154

The high R^2 value assumes a good predictive accuracy of the regression model, since 80.1% of the variance in the measure of ERP effort can be predicted by measuring “number of users”, “number of EDIs” and “number of conversions”.

But the average expected error (MMRE) of ±127% shows another picture. This seems to be a rather unsatisfactory result. In any case, regression analysis outperforms the ANGEL approach.

After finishing this step, Myrtveit & Stensrud (1999b) tested ANGEL and regression analysis for their practicality. Therefore, they conducted a survey with experienced practioners in a three-step experimental design.

Each participant was to estimate the effort on the basis of the initially suggested project characteristics, which are “users”, “sites”, “plants”, “companies”, “interfaces”, “EDIs”, “conversions”, “modifications”, “reports” and “number of modules”, and their value parameters in a first step. In a second step, the participants received the same information as in step one, and additionally the costs of the ten closest results identified by ANGEL. In the last step, the participants received the same information as in step one, plus the results of the regression analysis and the model itself.
Then, Myrtveit & Stensrud analysed which step yielded the best results. The experiment could not significantly increase the initial results. However, having an MMRE of 126%, the third step yielded the best results.

Widmer (2004) is the second author using multiple regression for ERP cost estimation. He aimed to develop a formula for predicting the total ERP costs.
In order to identify the “right” cost drivers, he conducted interviews with experts in a first step. Their suggestions were used as questions in a survey, and Widmer managed to receive valid replies of 42 Swiss companies. The analysis of the data set allowed him to identify number of users and project complexity as the two main cost drivers during an ERP lifecycle.

Based on empirical results and a regression analysis, Widmer presented a kind of formula for calculating ERP total costs in CHF for companies having between 25 and 100 employees:

$$ExpectedTotalCosts[CHF] = 160.000 + 9.500 \times Users + 56.000 \times Teamsize$$

With this formula, he attains an MMRE of 40% which can be evaluated as a good result.

Widmer calculates the effort for a time period of eight years and assumes to cover the whole lifecycle.
He argued that cost estimation of the whole lifecycle is more accurate than just considering the implementation costs. He reasons that with the argument that ERP vendors have different price policies. Some vendors have low initial costs but require a large amount of maintenance. Considering only the implementation costs would not capture these subse-

quently arising costs. This is a very important issue which will be discussed in more detail in chapter 4.

The third study using the regression method is already mentioned in earlier sections, like in the chapter about analogy-based estimation approaches, social-choice theory and DEA: Koch & Mitlöhner (2010) compared and assessed different existing effort estimation approaches empirically. Regression was one of them.
In their analysis, they regressed the following five variables that have an influence on ERP costs: number of users, interfaces, locations, modules, as well as type of ERP system.

For empirical validation of the different estimation approaches, they conducted a survey which yielded a data set of 39 valid ERP projects implemented in Australian companies.
In order to assess the results of the different estimation techniques, they used MMRE, MdMRE and pred (0.2) as quality criteria. The results are presented in the table below:

Table 3.17: Comparative results by Koch & Mitlöhner (2010, p. 275)

Method	Pred (0.2)	MMRE in %	MdMRE in %
Social Choice (weighted, Copeland)	0.39	543	50
Social Choice (weighted, Borda)	0.33	958	62
DEA	0.31	155	48
Analogy-based Estimation	0.23	48	43
Multiple Regression	0.05	1159	872

Regarding the MMRE, multiple regression yields the lowest accuracy and is the most unsuitable method. Analogy-based estimation shows the best results.
According to Koch & Mitlöhner (p. 276), this “underlines that social choice and DEA-based approaches are valid tools in this context”.

PAPERS CONTRIBUTING ONLY THEORETICAL DISCUSSION
Even the paper discussing different ERP estimation approaches is already stated in earlier sections, like in the chapter about analogy-based estimation approaches:

Stensrud (2001) compared and evaluated different existing effort estimation approaches theoretically. One approach in his comparative study was regression analysis.
In his paper, he emphasises the advantages of statistical techniques compared to non-parametric approaches. According to him (p. 421), only this approach deals "properly with stochastic variation, and therefore, the effort predictions from non-parametric approaches inspire less confidence […] than predictions from parametric approaches such as regression analysis that create expected values and provide confidence levels".
This argument and the flexibility of this method led him to conclude that regression analysis is the only parametric approach that makes "completely good sense used as a prediction system for ERP projects" (p. 422). Nevertheless, he recommended using the regression and analogy-based system ANGEL complementarily since it might add additional value.

3.3.7.2 CRITICAL APPRAISAL

Correlation and regression analysis are the most often applied approach within the topic of ERP effort estimation. The identified papers could be grouped into "correlation analysis", "regression analysis" and "theoretical discussion".
Since the single papers within these groups show similar strengths and limitations because the same research method is applied, each group will be critically discussed and assessed instead of regarding every paper.

In general, it can be said that all correlation and regression studies made different suggestions on how to measure the extent of an ERP project. Most of them attempt to map the extent by identifying suitable cost drivers. Table 3.18 shows the validated cost drivers of the mentioned papers.

Table 3.18: Validated cost drivers in correlation & regression analysis

	(Myrtveit & Stensrud, 1999)	(Widmer, 2004)	(Koch & Mitlöhner, 2010)	(Francalanci, 2001)	(Equey et al., 2008)
No. of users	x	x	x	x	
No. of EDI	x				
No. of data conversions	x				
Type of ERP system			x		
No. of modules			x	x	x
Per-module no. of submodules				x	
No. of locations			x		
No. of interfaces			x		
Team members		x			
Annual revenue				x	x
Enterprise is subsidiary					x
Management involvement					x
Ratio external to internal					x
Consultants' level of experience					x
Employee involvement					x
Procurement module					x
Production module					x
Sales/CRM module					x
Inventory module					x
Total number of employees				x	
Per-module number of users				x	

The table emphasises the different distributions of the validated cost driver candidates. They differ strongly. While some studies, like one by Myrtveit & Stensrud (1999), validated only technical aspects for defining ERP effort, other papers found organisational or human influences. Every single study showed interesting aspects, but verified only a small set of predictor variables.

The testing of different cost driver candidates implies that no generally accepted set of variables was found which can reliably measure the actual effort.

However, the different groups "correlation analysis", "regression analysis" and "theoretical discussion" will be discussed in the following.

CORRELATION ANALYSIS

Francalanci (2001) suggested considering the following six cost drivers for estimating the human resource effort in ERP implementations: “number of modules”, “number of sub-modules”, “revenue”, “total number of employees”, “total number of users” and “number of licences”. She was able to empirically prove her suggested cost drivers based on a data set of 43 SAP/ R3 projects.

Equey et al. (2008) confirmed the cost drivers “revenue” and “number of modules” but also found contradictory results. Furthermore, they did not regard Francalanci’s suggested variables “number of submodules” and “licence” but assumed new ones. The table below emphasises the similarities and differences of their studies.

Table 3.19: Similarities and differences in findings by Francalanci & Equey et al.

Finding	Cost Drivers
Confirmed	Revenue Number of modules
Contradictory finding	Number of users Number of employees
Not regarded	Number of submodules Licence
New	Enterprise subsidiary Management involvement ERP consultant’s level of experience Employee involvement Ratio of external consultants to internal employees SCM module Production module Sales /CRM module Inventory module

The different results might arise from the fact that Francalanci (2001) regarded the costs “human resources effort” whilst Equey et al. (2008) stated that they included hardware, external and licence in their costs. This seems to be a bit contradictory to the cost driver candidates which appear to not map these issues. But this is speculation.

Moreover, the identified relationships can just be assessed as low or slight strengths. None of the suggested show strong or high dependences.

In the context of this thesis, the main limitation of correlation analysis is seen in the fact that it does not provide a tool which enables a determination of the effort of future ERP projects.
Of course it gives a first indication of the critical factors or cost drivers responsible for incurring costs, but these do not allow prediction.

REGRESSION ANALYSIS
As in correlation and analogy-based analysis, the main question for regression analysis is which cost drivers are the “right” variables for predicting the ERP effort. The three papers identified arrived at different results in both, the relevant cost driver variables and the quality or suitability of regression models.

Stensrud & Myrtveit (1999) found empirical evidence for the following three cost drivers: “number of users”, “number of EDI” and “number of data conversions”. Their regression achieved an MMRE of 127%. Widmer (2004) confirmed the variable “number of users” but could only find evidence for one more variable, which is the number of “team members”. Attaining an MMRE of 40%, he yielded the best model quality. With five cost drivers, Koch & Mitlöhner (2010) regarded the most variables. But this did not produce a better result: presenting an MMRE of 1159%, this model showed the poorest quality.
The table below summarises the verified variables and presents the MMRE of the identified studies using regression analysis for ERP error estimation.

Table 3.20: Variables validated in regression analysis and model MMRE

	(Myrtveit & Stensrud, 1999)	(Widmer, 2004)	(Koch & Mitlöhner, 2010)
No. of users	x	X	x
No. of EDI	x		
No. of data conversions	x		
Type of ERP system			x
No. of modules			x
No. of locations			x
No. of interfaces			x
Team members		X	
MMRE	**127%**	**40%**	**1159%**

The MMRE shows the average prediction error which is a good quality indicator. The MMRE of 1159% presented by Koch & Mitlöhner is very high. This allows only one interpretation: the regression analysis in their model is not suitable to make reliable predictions of ERP effort estimation.
Indeed, the second highest value of 127% is remarkably below the first mentioned value, but still does not provide a satisfying result. The best model quality is achieved by Widmer with results of 40%. That means that the actual values deviate on average about 40% from the estimated ones. Admittedly, at first sight, this is still a big difference, but studies report cost overruns of 178% (Buckhout, Frey and Nemec: (1999). So a mismatch of 40% appears to provide progress.

However, regarding his identified variables, the good quality criteria seem to be rather unexpected at first sight. The variables “number of users” and number of “team members” do not seem to indicate the full extent of the ERP project, since the range of functions is not regarded. A good MMRE was more expected for the study of Koch & Mitlöhner who validated parameters like “number of modules” and “number of interfaces”.

Maybe these unexpected results are influenced by the different data sets. Widmer’s data set includes 42 companies having between 25 and 100 employees.

The data set of Koch & Mitlöhner consists of 39 companies having 1 to more than 2000 employees. Both studies use cases with different ERP vendors and are not limited to one supplier. Myrtveit and Stensrud (1999), in contrast, regarded just SAP/R3 implementations and included 48 cases with 7 to 2000 users, which indicates a higher employee quantity. Considering a smaller employee structure might be one reason why Widmer got the best quality result.

PAPERS CONTRIBUTING ONLY THEORETICAL DISCUSSION

Stensrud (2001) concluded in his comparative discussion that regression analysis is a very good technique to achieve proper results in ERP effort estimation. He reasons that with a proper dealing of variance, the provision of a numeric relationship and the flexibility of this technique, which does not have pre-defined input parameters. This links to the discussion what the "right" cost drivers are.

The author of this thesis follows Stensrud's arguments. Furthermore, when comparing the MMRE value of the different approaches, as shown in table 3.21, regression analysis yielded the best results, followed by analogy-based ANGEL.

Table 3.21: Comparative presentation of MMRE identified in different approaches

Method	**Author**	**MMRE in %**
Social Choice (Copeland)	(Koch & Mitlöhner, 2010)	543
Social Choice (Borda)	(Koch & Mitlöhner, 2010)	958
Multiple Regression Analysis	(Koch & Mitlöhner, 2010)	1159
	(Widmer, 2004)	40
	(Myrtveit & Stensrud, 1999)	127
Case-based Reasoning (ANGEL)	(Koch & Mitlöhner, 2010)	48
	(Myrtveit & Stensrud, 1999)	154
DEA	(Koch & Mitlöhner, 2010)	155

The author of this thesis comes to the conclusion that regression analysis seems to be one of the most promising approaches within ERP effort estimation.

3.3.8 CONCLUSION

This chapter aimed to find out which ERP cost estimation models exist and if they are suitable for providing reliable results. Therefore, a systematic literature review was conducted.

It found that the issue of ERP effort estimation is rather an unresearched issue which has only been scarcely discussed. Merely 14 papers dealt with this research field in the time period from 1997 to 2010. The seven identified approaches to move forward are very versatile, but none of them yield reliable, accurate outcomes. There is no generally accepted approach or model which is able to predict the costs of an ERP project. The most promising approach, showing the best MMRE, is the regression model developed by Widmer (2004).

This chapter presented the following seven identified approaches and discussed their suitability, strength and limitations:

- Transferability of estimation models for custom-built software (COCOMO II and COCOTS)
- Analogy-based models / Case-based reasoning
- Social-choice theory
- Data Envelopment Analysis
- Activity-based ERP effort estimation
- Organisational integration
- Correlation & regression analysis

The two approaches activity-based effort estimation and organisational integration do not provide a tool or technique for actually predicting the costs of future ERP projects. They might provide valuable knowledge but cannot be used directly in this context. Consequently, both approaches are excluded from the list of potentially suitable ERP effort prediction tools.

Furthermore, neither COCOMO II nor COCOTS were assessed as being useful prediction tools for ERP project prediction. COCOMO II was developed to predict the effort of custom-built software and consequently focuses mainly on the average productivity of the involved programmers and the project size, which is measured in source lines of codes (SLOC). They do not fit into the reality of ERP projects. Moreover, most of the suggested technical effort multipliers would hardly seem transferable to ERP since they do not map the prominent characteristics: They are packaged software, and the major challenge is integrating the pre-defined elements into the business processes of a specific organisation. The programming efforts play just a minor role in this context.
COCOTS estimates the effort for integrating COTS elements into the software system. Although this model regards some important aspects, it completely disregards others. Aspects like “user”, “interfaces” and “data conversion” are absolutely ignored.
The author of this thesis follows Stensrud’s and Francalanci’s opinion that these models are not suitable to be transferred to ERP projects and discards them as potential estimation approaches.

The four remaining approaches case-based reasoning, social-choice theory and data envelopment analysis can be assessed by comparing the quality criterion MMRE of the already conducted empirical researches. The results were presented in table 3.20 in the previous section.

Based on the MMRE, case-based reasoning (ANGEL) and multiple regression are evaluated as the most promising approaches.

Case-based reasoning is a non-parametric approach that aims to discover similarities between a future ERP project and completed projects in order to predict the result by using analogy-based tools, such as ANGEL.
The main question is what the right aspects to compare are. Myrtveit & Stensrud (1999) and Koch & Mitlöhner (2010) made different assumptions in this regard. Table 3.22 emphasises the different regarded aspects:

Table 3.22: Aspects concerned in case-based reasoning approaches

Project Characteristic	(Myrtveit & Stensrud, 1999)	(Koch & Mitlöhner, 2010)
No. of users	x	x
No. of sites /locations	x	x
No. of plants	x	
No. of companies	x	
No. of interfaces	x	x
No. of EDIs	x	
No. of conversions	x	
No. of modifications	x	x
No. of reports	x	
No. of modules	x	x

The main weakness of the approach is the non-parametric nature, which is not able to provide relationships between the project size and its effort. Because of that, Stensrud argues that regression analysis is the only parametric approach that makes “completely good sense used as a prediction system for ERP projects” (p. 422). According to him (p. 421), only this approach deals “properly with stochastic variation, and therefore, the effort predictions from non-parametric approaches inspire less confidence […] than predictions from parametric approaches such as regression analysis that create expected values and provide confidence levels”. This seems to be supported by the fact that the regression model of Widmer shows the best MMRE with a value of 40%.

Despite presenting the best MMRE results, the research of an accurate regression model is still in its fledgling stage.

The literature review identified just three papers applying this technique. Similar to case-based reasoning, the main question in developing a regression model is which cost drivers are suitable for mapping the extent of an ERP project. The three studies made different validations in this regard, which are summarised in the following table:

Table 3.23: Variables validated in regression analysis and model MMRE

	(Myrtveit & Stensrud, 1999)	(Widmer, 2004)	(Koch & Mitlöhner, 2010)
No. of users	x	X	x
No. of EDIs	x		
No. of data conversions	x		
Type of ERP system			x
No. of modules			x
No. of locations			x
No. of interfaces			x
Team members		X	

Generally, all regression studies made different suggestions how the extent of an ERP project can be measured. The table emphasises the different distribution of the validated cost driver candidates. They differ strongly. While some studies, like one by Myrtveit & Stensrud (1999), verified only technical aspects for defining ERP effort, other papers found organisational influences. Every single study shows interesting aspects but validated another set of predictor variables. The only uncontroversial variable is “number of users” which is verified by all of the three researches.

The different approaches show that there is no generally accepted set of variables that can reliably measure the actual effort.

However, regarding the identified cost driver variables, the good quality criteria attained by Widmer seems to be rather unexpected at first sight.

The variables “number of users” and number of “team members” does not seem to indicate the ERP project to its full extent, since the range of functions is not regarded. A good MMRE was more expected for the study by Koch & Mitlöhner who validated parameters like “number of modules” and “number of interfaces”.

One reason for this result might be that Widmer regards an essentially smaller employee structure in his data set than Koch & Mitlöhner. But this is speculation.

An extension of the regression model, i.e. regarding more cost driver candidates, seems to be a valuable approach for either confirming or broadening the findings of the previous studies.

However, besides giving an overview of existing models and approaches in ERP effort estimation, this chapter also touched the topic of defining costs. Francalanci, for example, regarded costs as “human resource effort”. Widmer, in contrast, regarded the ERP project for a period of eight years. According to him, this is the average duration of the ERP lifecycle. Others, like Equey et al. (2008), include hardware, external, and licence in their definition of cost, and others do not state their definition at all.
The next chapter addresses this issue and provides an overview of the different costs incurred and the aspects of costs in ERP projects.

3.4 SLR-RQ3: WHAT EFFORT TYPES ARE CONSIDERED IN ERP EFFORT ESTIMATION PAPERS?

The sub-chapter above presented different models and approaches for estimating the cost of ERP systems. Now this sub-chapter will focus on the costs themselves.

Since none of the presented cost estimation approaches are able to predict the ERP costs to their full extent, a deeper understanding of the general word “costs” seems to be necessary. Which aspects are considered under the category of costs?

The systematic literature review identified ERP costs as rather poorly defined and found a totally divergent research understanding of costs. Nearly all identified literature focused on different aspects or did not clarify their understanding of costs at all. There is a very fragmented, non-consensual view of ERP costs. No one has ever mapped an overall picture of all incurred ERP costs before within the research topic of ERP effort estimation.

In order to make reliable predictions, it is necessary to regard every single aspect that causes costs. Disregarding some of them might lead to inaccurate prognoses.
However, the literature review found non-transparency and fragmentation with regard to that issue in the 18 identified papers. An analysis of the papers showed that the term “effort” and “cost” are used synonymously, but there is a different understanding in the scope of the effort and cost considered. Furthermore, 11 of the papers do not clearly state

what is meant when using the term "effort" or "cost" even when these papers deal with models for effort estimation or analysing cost drivers. Table 3.24 shows the different effort scopes in the relation to the paper. In order to synthesise the effort types, the original wording had to be changed to a standard. Therefore, the naming of the cost types comes from the author of this thesis according to the most appropriate possible synthesis of the stated cost understandings in the research papers.

An analysis of all papers identified the following cost types: internal personnel costs, external personnel costs, organisational change costs, ERP software costs, license costs, business process reengineering costs, organisational change costs.

The most common understanding of effort lies in the personnel costs, which was stated in 6 papers. Thereof 2 papers consider internal and external personnel costs as the only effort type, whereas the other 4 papers considered personnel costs as just one part of the overall effort. The most detailed view of effort was found in the works of Kusters, Heemstra, & Jonker, (2007), Widmer, (2004), Koch & Mitlöhner (2010), as shown in table 3.7. These results demonstrate that there is not a common understanding in the scope of the term 'effort' or 'cost' in the context of ERP implementation. With the exception of the papers by Francalanci (2001) and Myrtveit & Stensrud (1999) considering only internal and external personnel effort, all other papers are dealing with a differing scope of effort types.

Table 3.24: Identified cost types in literature review

Paper	Internal Personnel Costs	External Personnel Costs	Hardware Costs	Licence Costs	ERP Software Costs	Organisational Change Costs	Business Re-engineering Costs	Not Stated
(Arb, 1997)								x
(Equey, Kusters, Varone, & Montandon, 2008)		x	x	X				
(Francalanci, 2001)	X	x						
(Kusters, Heemstra, & Jonker, 2007)								x
(Barki & Pinsonneault, 2002)			x		x	x	x	
(Stensrud, 2001)								x
(Widmer, 2004)		x	x	x	x			
(Koch & Mitlöhner, 2010)		x	x		x			
(Janssens, et al., 2008)								x
(Daneva, 2008b)								x
(Daneva, 2008c)								x
(Daneva & Wieringa, 2008)								x
(Daneva, 2007)								x
(Daneva, et al., 2008)								x
(Hansen, 2006)								x
(Barki & Pinsonneault, 2005)								x
(Kwon & Shin, 2005)	X	x						
(Myrtveit & Stensrud, 1999b)	X	x						
Frequency	3	6	4	2	3	1	1	11

Table 3.24 shows that 11 researchers do not state at all which costs they are writing about. It is not clear for the reader if they mean costs in general or which aspects are considered in their research.
In order to show how each single cost type and its covering elements are understood in this thesis, an explanation is given in the following section.

Internal personnel costs

Internal personnel costs are the total internal personnel costs in an organisation, which include elements like:

- Time and financial expenditures for visiting ERP fairs for evaluation of an ERP system
- Efforts for the allocation of internal personnel, such as key user team and IT department, with regard to the evaluation, implementation, and the maintenance of an ERP system
- Internal personnel effort for participating in analysis workshops and key user trainings
- Internal personnel efforts for maintaining the ERP system

External personnel costs

External personnel costs are different efforts for employment of external personnel, such as consulting costs. Examples of external personnel costs are:

- Consulting costs for the evaluation of an appropriate ERP system
- ERP consulting costs for installation work and workshops
- Consulting costs for business process analysis & tailoring of the ERP system
- Consultant costs for the realisation of key and/or end user trainings
- Project management costs

ERP software costs

Price policies of ERP vendors are quite different which makes a comparison difficult.
Fundamentally ERP software costs are defined as costs which are directly related with the ERP system and occur only once in this thesis.
ERP software costs occurring periodically are defined as licence costs presented further below. In general, ERP software costs include:

- Lump efforts for the purchase of an ERP system
- Training costs (insofar as these efforts are not already considered within the external personnel costs)

- Costs for reprogramming, insofar as this programming work is done externally. If programming work is done in-house by the company's personnel, the efforts are a matter of internal personnel costs.

Licence costs

Licence costs are software costs incurred on a periodic basis.

Hardware costs

Hardware costs are investment efforts for IT infrastructure and hardware, which are necessary for enabling the implementation and operation of the ERP system. This includes not only the initial investment but also the investment during the maintenance phase.

Business process re-engineering costs

These emerge through the attempt to increase the level of organisational integration. These are efforts for the analysis and conception of business processes.

Organisational change costs

Organisational change costs are a result of business re-engineering. These are the costs occurring due to the realisation of business re-engineering efforts, until an organisation has changed to the desired level of integration.

After the different cost types of ERP systems have been explored and summarised in order to map a picture of all ERP cost types, the next subsection analyses the influencing cost factors.

3.5 SLR-RQ4: WHAT ARE THE COST DRIVERS INFLUENCING ERP EFFORT?

Within the literature, many effort multipliers, aspects or cost drivers are discussed. The different effort estimation models presented in section 3.2 have already given an idea of them. In this thesis they are called "cost drivers".

As touched upon in section 3.2, some were already tested in empirical studies; others originated from theoretical frameworks and are not empirically confirmed yet. A lot of assumptions and speculations appear within the cost estimation of ERP systems.

To determine ERP costs, it is important to identify the influencing factors. Regarding this issue, the literature shows a quite similar picture as presented within the section cost types: Research does not provide an overall picture of these influencing factors.
Almost every researcher focused on single factors, but they were not synthesised into one concept before. Kusters, Heemstra, & Jonker stated that "no single comprehensive reference was found that presents a structured overview of cost drivers that affect the amount of costs required" (2007, p. 104).
Also Koch & Mitlöhner (2010, p. 277) argued in their comparative study in 2010 "for most we do not yet know a complete list of cost drivers for ERP implementations […]".

This thesis aims to fill this gap.
It pools all factors discovered in order to provide a broader view of ERP costs. To this end, a systematic literature review was conducted which distilled all factors.
It identified 11 papers making suggestions about cost drivers. In total, 64 cost driver candidates were found. They are presented in Table 3.25.

Table 3.25 Identified cost drivers in literature review

Cost Drivers	(Arb, 1997)	(Equey, et al., 2008)	(Francalanci, 2001)	(Kusters, et al., 2007)	(Stensrud, 2001)	(Widmer, 2004)	(Koch & Mitlöhner, 2010)	(Daneva, 2010)	(Daneva et al., 2007, 2007b. 2008, 2008c)	(Hansen, 2006)	(Myrtveit & Stensrud, 1999b)	Freq.
No. of locations					x		x		x		x	7
No. of total users	x		x	x	x	x	x				x	7
Analyst capability (COCOMO II)				x					x			5
No. of modules			x	x	x		x				x	5
Personnel continuity (COCOMO II)				x					x			5
Programmer capability (COCOMO II)				x					x			5
Application experience (COCOMO II)									x			4
Database size (COCOMO II)									x			4
Documentation (COCOMO II)									x			4
Platform volatility (COCOMO II)									x			4
Product complexity (COCOMO II)									x			4
Program language and tool experience (COCOMO II)									x			4
Required implementa-tion schedule (COCOMO II)									x			4
REUSE (COCOMO II)									x			4
Use of software tools (COCOMO II)									x			4
No. of data conversions				x	x						x	3
No. of interfaces				x	x						x	3

No. of modifications				X			x				x	3
No. of reports				X	x						x	3
Ability to change				X						x		2
Consulting experience		x		X								2
No. of EDIs					x						x	2
No. of organisational units or departments				X	x							2
No. of submodules	x		x									2
Team maturity				X						x		2
Team quality				X						x		2
Willingness to change				X						x		2
Availability of management				X								1
Availabilty of business users				X								1
Commitment management				X								1
Complexity of interfaces				X								1
Complexity of business processes				X								1
Complexity of data				X								1
Complexity of reports				X								1
Complexity of transactions				X								1
Consultant knowledge				X								1
Consultant quality				X								1
Critical attitude of users				X								1
Employee involvement		x										1
Fit of ERP system and organisation				X								1
Infrastructure				X								1
Management				X								1
Management involvement		x										1
Maturity of processes				X								1
Maturity of technology				X								1
No. of transactions				X								1
No. of business processes				X								1

No. of companies											x	1
No. of total employees			x									1
No. of modified screens					x							1
No. of stakeholders				x								1
No. of user groups				x								1
No. of users per module			x									1
Quality of business users				x								1
Revenue			x									1
Stability of organisation				x								1
Steering management				x								1
Team composition				x								1
Test approach				x								1
Quality of tools				x								1
Type of module		x										1
User quality				x								1
Vision				x								1
Ratio external / internal		x										1

Table 3.25 shows who of the 11 authors mentioned which cost drivers. Furthermore, the frequency of statement of every cost driver is presented.

With a frequency of seven times, the most often mentioned cost drivers are “number of company locations” and “number of total users”. “Analyst capability”, “number of modules”, “personnel continuity” and “programmer capability” are stated five times.

It is interesting to note that 39 of the 64 cost drivers were not found more often than once. Thus most of the identified cost drivers were only stated once.

However, it must be said that Daneva’s four papers have a strong impact on the frequencies because she stated the same cost drivers taken from the COCOMO II method in all of her studies.

Just one of the stated cost drivers is already empirically confirmed. It is the “number of users” which was verified by the studies of Myrtveit & Stensrud (1999), Widmer (2001) and Koch & Mitlöhner (2010).

Other cost drivers, like “number of team members”, “number of locations”, “number of modules” and “number of interfaces”, were only validated once, as identified in chapter 3. Some cost drivers were only mentioned on a hypothetical basis.

This comprehensive overview of the stated cost driver candidates seems to be a good starting point for empirical validation.

3.6 SLR-RQ5: WHAT PROJECT OR LIFECYCLE PHASE IS CONSIDERED IN ERP EFFORT ESTIMATION PAPERS?

Besides identifying the different cost types and the cost-influencing cost drivers, another important aspect for cost estimation is to analyse at what time stages costs arise and to define the time period of an ERP lifecycle.
This section will present different ERP lifecycle models with their individual lifecycle phases at first, before analysing the considered timespan of previous researches and models.

3.6.1 MODEL PRESENTATION

ERP software usually goes through different phases in its lifetime. The divergent suggestions about classifying these stages will be presented in the following. Since several researchers have developed lifecycle models, exemplary ones are presented in the following, structured according to the number of phases they suggest for characterising the lifecycle.

THREE STAGES
Chang, Gable, Smythe & Timbrell (2000), for example, argue that the traditional IS approach to describe lifecycles in terms of development, implementation and maintenance is not appropriate for ERP, since it conceals the repetitive character of ERP. They suggest dividing the lifecycle into the three stages pre-implementation, implementation and post-implementation, and emphasise that these stages “continue throughout the lifetime of the

ERP as it evolves with the organisation" (Chang, et al., 2000, p. 494). According to Stefanou (2001, p. 208), the three different phases include the following activities:

I. Pre-implementation: Requirements definition, business case and software selection

II. Implementation: Bug analysis, tailoring, and project and change management

III. Post-implementation: Rollout, upgrades and payback review.

FOUR STAGES

Authors like Markus & Tanis (2000) and Stefanou (2001) categorised the ERP lifecycle into four phases. Markus & Tanis (2000) emphasise the 'post-implementation' phase of the ERP lifecycle and divide this stage into the two categories 'Shakedown' and 'Onward & Upward'. In contrast, Stefanou focuses more on the phase before ERP is implemented. He regards this stage 'pre-implementation' in more detail by differentiating it into 'Business Vision' and 'ERP-Selection'. Their suggestions are considered below:

Markus, et al. (2000) proposes the division presented in the following four stages:

I. Chartering Project: This phase includes organisational decisions about the project's aims, its integration process and its budget, and the definition of the business case.

II. Project (Configure & Rollout): This phase means to activate ERP and to configure the system until it is rolled out.

III. Shakedown: At this stage, organisations become familiar with the new ERP, and its utilisation becomes routine for the applicators.

IV. Onward & Upward: Here, ERP is replaced through an update or another system.

Stefanou (2001) highlights the importance of the ex-ante ERP evaluation in his four-stage model. According to him, the lifecycle can be assessed by regarding the following phases.

I. Business Vision: Defining and figuring the long-term view of an organisation.

II. ERP Selection: Examining and determining the organisational requirements, analysing the capabilities and limitations of business processes, and the organisational willingness to change. Selecting the vendor and the system modules according to the findings.

III. ERP Implementation: Estimating the costs and benefits needed for ERP integration.

IV. ERP Operation | Maintenance | Evolution: Evaluating the subject involved in operations, maintenance and evolution.

The four stages are affected by an enduring evaluation of costs, benefits and risks, including tangible and intangible measurements for both operational performance and strategic positioning of an organisation. Assuming repetitions of some activities within his model, Stefanou concludes his four steps are not purely sequential.

FIVE STAGES

A five-stage approach for describing an ERP lifecycle was developed by Ross (1999). Although the category names differ from the four-step approach by Markus & Tanis (2000), her model demonstrates a certain textual similarity to it. While the first four phases can be regarded as rather equivalent, the difference lies in the fifth stage.

I. Design: Refers to managerial decisions whether the best 'practice principles' embedding ERP can be adapted to the organisation and the extent to which business processes become standardised.

II. Implementation: Means getting the system and the user up and running.

III. Stabilisation: Identifies cleaning up data and parameters in order to adjust to the new environment and bug solving in the software as major activities.

IV. Continuous improvement: Means the extension of bolt-ons, the activation of new modules or the addition of other functional features.

V. Organisational transformation: At this stage, the organisational behaviour is analysed to determine if anything has transformed since ERP integration. Unfortunately, Ross has not found a company from her research which felt it had transformed itself.

SIX STAGES

Most project phases were presented in a very detailed six-stage model by Esteves & Pastor (1999). Like Stefanou (2001), the authors found two stages of importance before ERP implementation, but the respective fourth phases 'Continuous Improvement' presented by Ross (1999) and 'Onward & Upward' presented by Markus & Tanis (2000) are subdivided into two more categories, namely 'Evolution' and 'Retirement'.

The six stages of the ERP lifecycle developed by Esteves & Pastor (1999) are presented in the following:

I. Adoption & Decision: This is the phase in which the needs, requirements, objectives and benefits for moving to ERP are defined.

II. Acquisition: At this stage, selecting the system which is most suitable to the business processes and agreements about terms and conditions with the vendors are the main activities.

III. Implementation: This category refers to activating the system and tailoring the package.

IV. Use & Maintenance: This phase is the one during which the system is used, bugs are corrected and expected benefits are provided. In fact, it is very similar to the ‘Stabilisation’ or ‘Shakedown’ phases in the models by Ross (Ross, 1999) and Markus & Tanis (Markus & Tanis, 2000).

V. Evolution: Means the integration of further capabilities, like the extension of additional features or the external linkage to customers and suppliers.

VI. Retirement: At this final stage, ERP is substituted by an update or another system.

Finally, the six phases are enduringly affected by the four different dimensions “change management within the organisation’, ‘employees’ abilities to cope with ERP’, ‘the fit between ERP and organisational processes’ and the ‘functionality of the ERP product’ itself.

This section presented exemplary lifecycle models. Although they have different priorities and differentiate some phases more or less, a certain similarity can be observed in all presented models. However, this section clarifies that an ERP system goes through different stages in its lifetime and, of course, costs can be incurred in every single phase.
The next section will analyse if research has already identified specific costs in specific lifecycle phases.

3.6.2 CONSIDERED LIFECYCLE PHASES IN ERP EFFORT ESTIMATION APPROACHES

A holistic consideration spanning the whole ERP lifecycle seems to be essential to making reliable predictions about the total costs incurred by ERP systems. Otherwise, some costs might be disregarded which can result in strong cost overruns. The coverage of ERP costs can only be complete and reliable if all phases are considered.

A systematic literature review was conducted to discover how research to date has considered the costs during the different lifecycle phases.

It found 18 relevant papers, but identified that 11 of them concentrate just on the implementation phase or do not state the duration of the consideration at all. Table 3.26 gives an overview of the findings.

Table 3.26: Regarded lifecycle phases in literature

Paper	Evaluation Phase	Implementation Phase	Maintenance Phase	Not Stated
(Arb, 1997)		x		
(Equey, et al., 2008)		x	x	
(Francalanci, 2001)		x		
(Kusters, et al., 2007)		x		
(Barki & Pinsonneault, 2002)		x		
(Stensrud, 2001)				X
(Widmer, 2004)		x	x	
(Koch & Mitlöhner, 2010)		x		
(Janssens, et al., 2008)		x		
(Daneva, 2008b)				X
(Daneva, 2008c)				X
(Daneva & Wieringa, 2008)				X
(Daneva, 2007)				X
(Daneva, et al., 2008)				X
(Hansen, 2006)				X
(Barki & Pinsonneault, 2005)		x		
(Kwon & Shin, 2005)		x		
(Myrtveit & Stensrud, 1999b)		x		
Frequency	**0**	**11**	**2**	**7**

Most papers are suggestive of the view that costs arise only during the implementation phase. Only Widmer (2004) and Equey et al. (2008) offer a broader view of this issue. Widmer regards the implementation and maintenance phase of the ERP lifecycle in order to estimate the total costs more accurately.

He makes an important point when explaining that it is essential to regard more than the implementation phase simply because ERP vendors have different price policies in their consulting or licence cost models. This means that one vendor might have low implementation costs but high annual maintenance costs, or the other way around.

However, it is important to state that none of the found papers regard the evaluation phase in their cost estimation. This ERP lifecycle phase seems to be nearly unexplored to date.

3.6.3 CRITICAL APPRAISAL

Section 3.6.2 identified a research focus on the implementation phase of the ERP lifecycle. However, the evaluation phase and the maintenance phase are rather unexplored to date.

This issue raises some problems. It is questionable whether the analysed incurred costs during the implementation phase are suitable to give an accurate overall cost prediction.

As touched upon earlier, Widmer (2004) makes an important point when referring to the different price policies of the vendors, which might result in having low implementation costs but high annual maintenance costs, or vice versa.
Another point is that activities which are typically found in the maintenance phase are absolutely unregarded in the ERP cost estimation. Neither costs for tailoring or further development nor costs for business process re-engineering activities during later phases are considered (Barki & Pinsonneault, 2002).
The consideration of all costs actually incurred is unlikely when regarding just the implementation phase.

None of the identified researches deal with the question whether there are cost relationships between the lifecycle phases. Aside from the missing relationship of implementation and maintenance costs, there is also no discussion about the cost relationship of the ERP evaluation phase and other project or lifecycle phases.

It can be hypothesised that the selection process during the evaluation phase has a strong impact on the costs emerging in the implementation and maintenance phases.
If a company spends less effort in the ERP selection process, for example, it would not be surprising that there are several misfits between the organisational workflow and the ERP functions. This might lead to high tailoring costs during the implementation and maintenance phases.

However, the author of this thesis identifies a lack of cost research during the evaluation and maintenance phases of the ERP lifecycle.
In order to make adequate predictions about the total ERP costs, it might be necessary to provide a more detailed view of each single phase, the sum of which then enables the mapping of an overall picture of the costs incurred during the whole ERP lifespan.

3.7. CONCLUSION

This section identified that ERP costs are a quite undefined issue. Most researchers do not state their definition of costs within relevant ERP effort estimation papers, and the researchers who state which costs they consider do not have a consensual understanding of ERP costs. Every researcher focuses on a different aspect and no one has ever mapped an overall picture of all incurred ERP costs before. All found literature presents just extracts. That is why the author of this thesis has grouped the cost aspects into the three dimensions “cost type”, “cost drivers” and “ERP lifecycle phases”.

Firstly, seven different cost types could be identified.
Secondly, this thesis fills the gap of a lacking comprehensive reference that presents a structured overview of cost drivers by pooling all cost driver candidates ever stated.
Thirdly, it identifies a disregard of important lifecycle studies in previous research. The findings are presented in more detail below:

I.
In total, the systematic literature review found seven different cost types which are “internal personnel costs”, “external personnel costs”, “ERP software costs”, “licence costs”, ”hardware costs”, “organisational change costs”, and “business process re-engineering costs”.

II.
Filling the gap of a lacking comprehensive list by pooling 64 different cost driver candidates from different studies and research work. These are as follows:

- No. of locations
- No. of modules
- Database size
- Product complexity
- REUSE
- No. of interfaces
- Ability to change
- No. of submodules
- Team quality
- Availability of users
- Complexity of data
- Consultant knowledge
- Critical attitude of users
- Infrastructure
- Maturity of processes
- No. of business processes
- No. of modified screens
- No. of users per module
- Stability of organisation
- Test approach
- User quality
- No. of total users
- Personnel continuity
- Documentation
- Program language
- Use of software tools
- No. of modifications
- Consulting experience
- Team maturity
- Willingness to change
- Commitment management
- Complexity of reports
- Consultant quality
- Employee involvement
- Management
- Maturity technology
- No. of companies
- No. of stakeholders
- Quality business users
- Steering management
- Quality of tools
- Vision
- Analyst capability
- Application experience
- Platform volatility
- Req. implement. schedule
- No. of data conversions
- No. of reports
- No. of EDIs
- No. of depts. or units
- Management availability
- Complexity of interfaces
- Complexity of bus. processes
- Complexity of transactions
- Fit of ERP system & organ.
- Management involvement
- No. of transactions
- No. of total employees
- No. of user groups
- Revenue
- Team composition
- Type of module
- Ratio external to internal

Besides identifying the different cost types and the cost-influencing cost driver candidates, another important aspect of cost estimation is to analyse at what time stages costs occur.

III.

Identification of lacking research of whole lifecycle

Since an ERP system typically goes through different life phases in its lifetime, within this thesis the categories of the different model phases offered by the literature are analysed. It found that the concept of ERP lifecycles is rather similar, but differs in the quantity of stages used to describe the cycles. Their quantity varies from three to six.

In a further step, this thesis analyses the consideration of lifecycle phases in the context of incurred ERP cost estimation. The literature review found that mainly all research observes just the implementation phase. The incurring costs within other phases are completely ignored in most studies. Since the regard of maintenance and evaluation are important factors for the total costs, ignoring them could be fatal for ERP prediction concepts.

In summary, this chapter presents an unsatisfactory picture of cost understanding in ERP cost estimation approaches. A broader concept, which regards all the different cost types, all phases of the ERP lifecycle, and a specific mapping of all cost drivers suggested to date, might be a first step in making more precise predictions about ERP costs in the future. Within this thesis the author develops such a concept in the next chapter.

4. CONCEPTUAL FRAMEWORK

The findings of the literature review are salient for understanding this chapter. Therefore, they are briefly summarised. The previously presented chapters found the following main aspects:

I. No suitable model or prediction formula exists

There are many ERP prediction approaches; however, none of the identified models are able to predict ERP cost estimation to its full extent. With an MMRE of 40%, Widmer (2004) presented the best approach, but it also shows some questionable limitations.

II. Costs are an unexplored research issue

Costs in ERP cost estimation are not a well-researched issue. Many researchers do not state at all which costs they regard. It might be possible that some cost types are not considered, which may lead to inaccuracy. This thesis develops three dimensions of costs, which are "cost types", "cost drivers" and "timespan/lifecycle".

II.I Cost Types

This thesis found seven different cost types that had not been brought into one model before.

II.II Cost Drivers

The consideration of cost driver candidates is relatively fragmented. Almost every researcher proposes different ones. Since the disregard of all possible incurring costs is fateful for reliably predicting the overall costs, it is important to identify them all.

The literature review identified 64 cost drivers. They are pooled for the first time within this thesis.

II.III Timespan/Lifecycle Phases

Most research focuses just on the implementation phase and ignores the other stages of the ERP lifespan. This aspect might have fatal consequences for cost prediction since some of the arising costs could remain completely unregarded.

This thesis aims to synthesise the three dimensions “cost types”, “cost drivers” and “timespan/ lifecycle” into one research concept in order to contribute to a better understanding of costs and to develop a model for prediction the costs of an ERP project.

This chapter is structured to firstly present the rationale of the conceptual framework.
Then, it addresses the issue of cost types. This thesis selects the suitable cost types for the further research of this study and presents the exclusion criteria. Then, it selects and categorises the cost drivers. Next, an ERP model is chosen, which should be applied in this thesis, and selects the timespan for this study.
Next, it develops a research framework which conceptualises the three dimensions “cost types”, “cost drivers” and “timespan/ lifecycle” into one framework. Based on this concept, this thesis formulates three research questions which should access the issue. They are presented in chapter 4.6.

4.1 RATIONALE OF THE CONCEPTUAL FRAMEWORK

As mentioned before, this thesis aims to synthesise the three dimensions of costs into one conceptual model. The conceptualisation is based on the notion that the total costs of ERP are composed of the different cost types “internal personnel”, “external personnel”, “hardware”, “licence” and “ERP software” costs.
The cost drivers determine the extent of costs of each cost type. That means that cost drivers should be allocated to the cost types, not to the total costs.
Furthermore, it seems to be very important to regard the whole lifespan in order to provide accurate information about costs. Therefore, the cost types need to be assigned to the different lifecycle phases.

This differentiated consideration of the ERP costs enables a division of the total costs into smaller entities. This allows a better planning and controlling of ERP costs and permits a better understanding of the incurring costs of an ERP project.

This framework aims to develop a concept that includes the different cost types and cost drivers during the whole ERP lifecycle.

The next section regards and then selects the identified cost types suitable for the further research of this study.

4.2 COST TYPES

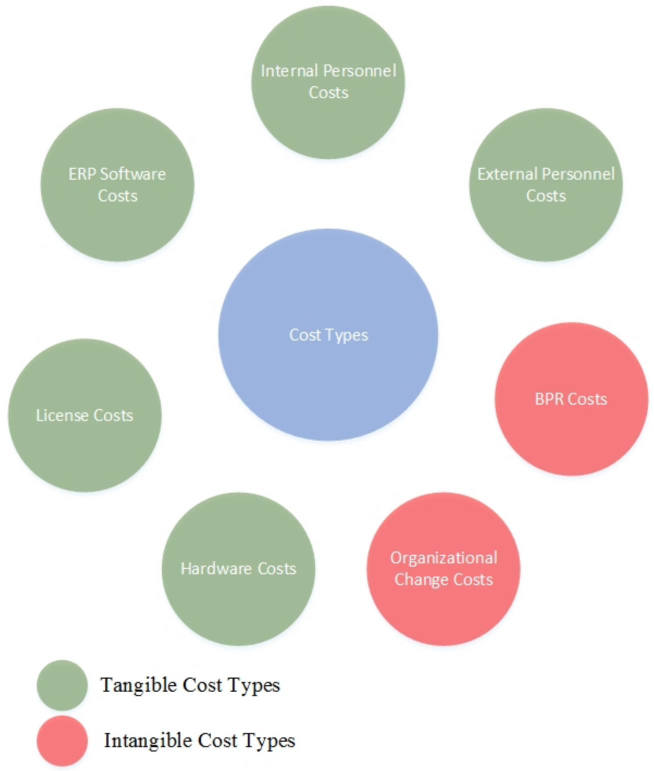
The first conceptualisation aspect in this thesis is the different cost types of ERP systems. The literature review identified the seven different cost types: “internal personnel costs”, “external personnel costs”, “hardware costs”, “licence costs”, “ERP software costs”, “organisational change costs” and “business process re-engineering costs”.

As mentioned above, the differentiation of the ERP costs enables the division of the total costs into smaller entities. This allows a better planning and control of ERP costs and permits a better understanding of the incurring costs of an ERP project.

4.2.1 SELECTION OF COST TYPES

On closer examination, a separation of the seven identified cost types into two groups seems to be ascertainable. One category is the tangible cost types and the other one is the intangible cost types. Figure 4.1 illustrates the categorisation of the identified cost types.

Figure 4.1: Tangible & intangible costs



Whilst tangible costs are rather easy to measure, intangible costs are much harder to quantify. Chapter 4.1 shows the understanding of these cost types. The defined activities are substantially an inherent part of either internal personnel costs or external personnel costs or both.

In consequence, the organisational change costs and business process re-engineering costs are not regarded as separate cost types in the further research. Their operations are part of the internal and external personnel costs. As a result, the two cost types are excluded.

4.2.2 RELEVANT COST TYPES IN THIS THESIS

The five tangible cost types "internal personnel costs", "external personnel costs", "hardware costs", "licence costs" and "ERP software costs" remain in this thesis and are used to develop the conceptual framework in chapter 4.5.

Although they are mainly presented in the previous chapter, they should each be briefly defined in order to clarify the understanding of them in this thesis.

1. Internal personnel costs

Full internal personnel costs can be associated with the ERP system within an organisation. Internal personnel costs include:

- Time and financial expenditures for visiting ERP fairs for evaluation of an ERP system
- Efforts for the allocation of internal personnel, such as key user team and IT department, with regard to the evaluation, implementation, and the maintenance of an ERP system
- Internal personnel efforts for participating in key user training and analysis workshops
- Internal personnel efforts for the maintenance of an ERP system
- Internal personnel efforts arising through business process re-engineering work
- Internal personnel efforts induced by the organisational change process (e.g. duplication of work)
- Project management efforts

2. External personnel costs

Several efforts for the employment of external personnel, such as consulting costs. Examples of external personnel costs are:

- Consulting costs for the evaluation of an appropriate ERP system
- ERP consulting costs for installation work and workshops
- Consulting costs for business process analysis, insofar as this is directly associated with the ERP system
- Consulting costs for the realisation of key user and/or end user training
- Project management efforts
- External personnel efforts arising through business process re-engineering work
- External personnel efforts induced by an organisational change process (e.g. duplication of work)

3. ERP software costs

Price policies of ERP vendors are quite different, which makes a comparison difficult. Fundamentally, ERP software costs are defined in this thesis as costs which are directly related to the ERP system and occur only once. ERP software costs occurring periodically are defined as licence costs presented further below. However, ERP software costs include:

- Lump efforts for the purchase of an ERP system
- Training costs
  (insofar as these efforts are not already considered within the external personnel costs)
- Costs for tailoring/ reprogramming, insofar as this programming work is done externally. If programming work is done in-house by the company personnel, the efforts are a matter of internal personnel costs.

4. Licence costs

Licence costs are software costs incurred on a periodical basis.

5. Hardware costs

Hardware costs are the investment efforts into IT infrastructure and hardware which are necessary for enabling the implementation and operation of the ERP system. This includes not only the initial investment, but also the investment during the maintenance phase.

4.3 COST DRIVERS

ERP cost drivers are the influencing variables that explain the emerging costs of ERP systems. Section 4.2 presented a comprehensive overview of cost driver candidates identified by the conducted systematic literature review. They are as follows:

Table 4.1: Overview of cost driver candidates

▪ No. of locations	▪ No. of total users	▪ Analyst capability
▪ No. of modules	▪ Personnel continuity	▪ Application experience
▪ Database size	▪ Documentation	▪ Platform volatility
▪ Product complexity	▪ Program language	▪ Required implem. schedule
▪ REUSE	▪ Use of software tools	▪ No. of data conversions
▪ No. of interfaces	▪ No. of modifications	▪ No. of reports
▪ Ability to change	▪ Consulting experience	▪ No. of EDIs
▪ No. of submodules	▪ Team maturity	▪ No. of depts. or units
▪ Team Quality	▪ Willingness to change	▪ Availability management
▪ Availability of users	▪ Commitment management	▪ Complexity of interfaces
▪ Complexity of data	▪ Complexity of reports	▪ Complexity of business processes
▪ Consultant knowledge	▪ Consultant quality	▪ Complexity of transactions
▪ Critical attitude of users	▪ Employee involvement	▪ Fit of ERP system & organisation
▪ Infrastructure	▪ Management	▪ Management involvement
▪ Maturity of processes	▪ Maturity of technology	▪ No. of transactions
▪ No. of business processes	▪ No. of companies	▪ No. of total employees
▪ No. of modified screens	▪ No. of stakeholders	▪ No. of user groups
▪ No. of user per module	▪ Quality of business users	▪ Revenue
▪ Stability of organisation	▪ Steering management	▪ Team composition
▪ Test approach	▪ Tools quality	▪ Type of module
▪ User quality	▪ Vision	▪ Ratio external to internal

These cost drivers should be considered for their inclusion or exclusion in the framework, and synthesised and structured in order to convey them into a framework.

4.3.1 SELECTION OF COST DRIVERS

There are four reasons for excluding cost driver candidates from the research concept of this study:
Firstly, this section aims to consider only ERP related cost drivers. Since chapter 3 identified the transferability of software development models to ERP cost estimation as not suitable, the cost drivers originating from those models, like COCOMO, are excluded.

The second reason is to avoid inaccuracy and confusion about the meaning of a cost driver candidate. For example, the variable “infrastructure” allows several different interpretations. In order to provide a unique understanding of aspects, these variables were either reformulated or excluded.

The third motivation for excluding cost drivers is to avoid confusion within the survey conducted later. Since this research is based on a web-based interview design, the questions have to be easy and self-explanatory in order to ensure comprehensibility. The candidate “number of business processes” was excluded because the definition of this variable would have been ambiguous.

The reduction of complexity is the fourth reason for excluding certain candidates. As touched upon before, this thesis is based on a web-based survey design and needs to ensure that all cost drivers can be enquired about without any uncomfortableness for the interviewee.

Table 4.2 gives an overview of the excluded cost driver candidates and the justification for such exclusion.

Table 4.2: Reasons for the omission of cost driver candidates

Cost Driver	**Reason for Omission**
Analyst capability	Software development related COCOMO II cost driver
Personnel continuity	Software development related COCOMO II cost driver
Programmer capability	Software development related COCOMO II cost driver
Application experience	Software development related COCOMO II cost driver
Database size	Software development related COCOMO II cost driver
Documentation	Software development related COCOMO II cost driver
Platform volatility	Software development related COCOMO II cost driver
Product complexity	Software development related COCOMO II cost driver
Program language & tool experience	Software development related COCOMO II cost driver
Required implementation schedule	Software development related COCOMO II cost driver
REUSE	Software development related COCOMO II cost driver
Use of software tools	Software development related COCOMO II cost driver
No. of data conversions	Not measurable
No. of submodules	Not measurable
Complexity of transactions	Inaccurate definition
Infrastructure	Inaccurate definition
Management	Inaccurate definition
Maturity of technology	Not measurable
No. of transactions	Not measurable
No. of business processes	Not measurable
No. of companies	Inaccurate definition
No. of modified screens	Not measurable
No. of stakeholders	Not measurable
No. of users per module	Not measurable
Steering management	Inaccurate definition
Test approach	Inaccurate definition
Quality of tools	Not measurable
Vision	Not measurable
Ratio external to internal	Not measurable
Total	**29**

In total, 29 cost driver candidates are excluded. The remaining 35 cost drivers are presented and structured in the next section.

4.3.2 RELEVANT COST DRIVERS IN THIS THESIS

The included 35 cost drivers are presented in table 4.3 below. In order to arrange them more clearly, they are structured according to Hansen's (2006) three-dimensional (organisation, technical, and situational) suggestions. The selected cost driver candidates have been allocated a variable name from x1-x35.

Table 4.3 Overview of included cost drivers

Organisational Cost Drivers	Technical Cost Drivers	Situational Cost Drivers
x1: No. of locations	x5: ERP system	x16: No. of consultants
x2: No. of organisational units or depts.	x6: No. of interfaces	x17: No. of project members
x3: No. of total employees	x7: No. of modifications	x18: Ratio external / internal
x4: Revenue	x8: No. of reports	x19: Fit ERP system / organisation
	x9: No. of EDIs	x20: Team quality
	x10: No. of total users	x21: Team maturity
	x11: No. of user groups	x22: Team composition
	x12: Type of modules	x23: Availability of management
	x13: Complexity of data	x24: Availability of business users
	x14: Complexity of interfaces	x25: Consulting experience
	x15: Complexity of reports	x26: Consulting quality
		x27: Critical attitude of users
		X28: User quality
		X29: Employee involvement
		x30: Management involvement
		x31: Maturity of processes
		x32: Complexity of business processes
		x33: Stability of organisation
		x34: Willingness to change
		x35: Motivation of implementation team

4.4 LIFECYCLE PHASES IN THIS THESIS

All identified studies mainly consider only the emerging costs during the implementation phase and ignore incurring costs during other phases of the lifecycle. In consequence, the arising costs during other stages may previously not have been analysed.

This thesis attempts to analyse the costs during each lifecycle phase in order to regard the long-term costs of ERP systems. To this end, the lifecycle phases first have to be defined for this thesis. Chapter 3.3 presented different models. Despite having different stages, they show a certain similarity.

For this thesis, a three-phase cycle seems to be the most adequate model. By grouping the ERP lifecycle into the categories “evaluation”, “implementation” and “maintenance”, all costs should be covered. A more differentiated view seems to bring no additional value.

A further important aspect is that this thesis is based on a survey design in which the respondents will be asked about the costs they incur at each stage of the lifecycle. Since it is very difficult to differentiate the single stages in practise (Markus, et al., 2000), the categorisation should be as simple, plain and comprehensible as possible in the survey design in order to minimise the drop-out rate.

The author of this thesis will basically follow the suggestion by Chang (Chang, et al., 2000) and distinguish between the three most basic stages presented in chapter 3.6.1.

Some modifications will be made in this thesis with regard to Chang’s suggested lifecycle.

Unlike Chang’s model, the phases will be called ‘evaluation’, ‘implementation’ and ‘maintenance’ to provide a more concrete idea of what happens in the individual phases.

Furthermore, ERP rollout is a component of implementation in this model and will not be regarded in the ‘maintenance’ category.

The timespan of the regarded costs is an important matter. This thesis assumes a 10-year lifecycle of an ERP system (D. Abts & Mülder, 2009)

To summarise, this framework regards the cycle phases as mentioned below:

I. Evaluation:

Definition of requirements, business case and ERP selection

II. Implementation:

Activation of ERP and system configurations (bug analysis, tailoring and project and change management) until rollout

III. Maintenance:

Retain the system in good condition, system improvement or extension and ERP updates and employee training.

IV. Lifespan:

The lifespan of an ERP project after its implementation is bound to be 10 years. This framework assumes that the maintenance phase will consequently last 10 years.

4.5 RESEARCH CONCEPT

One aim of the study is to analyse the cost structure of an ERP project.
Therefore, every cost type needs to be assigned to the lifecycle phases in order to provide a detailed overview of all incurring costs at all time stages. This provides a more holistic approach for analysing the long-term costs of ERP systems.

Combining the five different cost types and three different lifecycle phases will theoretically result in 15 different cost fields[14].
But, according to the definition given in the previous chapters, the cost types "hardware", "software" and "licence costs" cannot emerge during the evaluation phase.
Just internal and external personnel costs are bound to be incurred during this phase. Consequently, this framework regards twelve cost fields which are graphically presented in table 4.4.

Table 4.4 Overview of cost fields

	Internal Personnel Costs	External Personnel Costs	ERP Software Costs	Licence Costs	Hardware Costs
Evaluation	y1	y2	-	-	-
Implementation	y3	y4	y5	y6	y7
Maintenance	y8	y9	y10	y11	y12

[14] The variables emerging from the combination of cost types and lifecycle phases are named cost fields.

With this breakdown, ERP costs will be much more transparent than just considering ERP costs as one undefined cost pool. The undefined costs to date can be assigned to a certain cost field and, therewith, an ERP cost structure can be ascertained. The cost structure is able to averagely state the distribution of costs to each single cost field.

The different twelve cost fields will be analysed for their incurring costs. Therefore, they are regarded as dependent variables and examined for their relationship with the 35 cost drivers.

Figure 4.2 shows the conceptual framework in a graphical presentation. As described, the conceptual framework consists of the three dimensions lifecycle phase, cost types, and cost drivers.

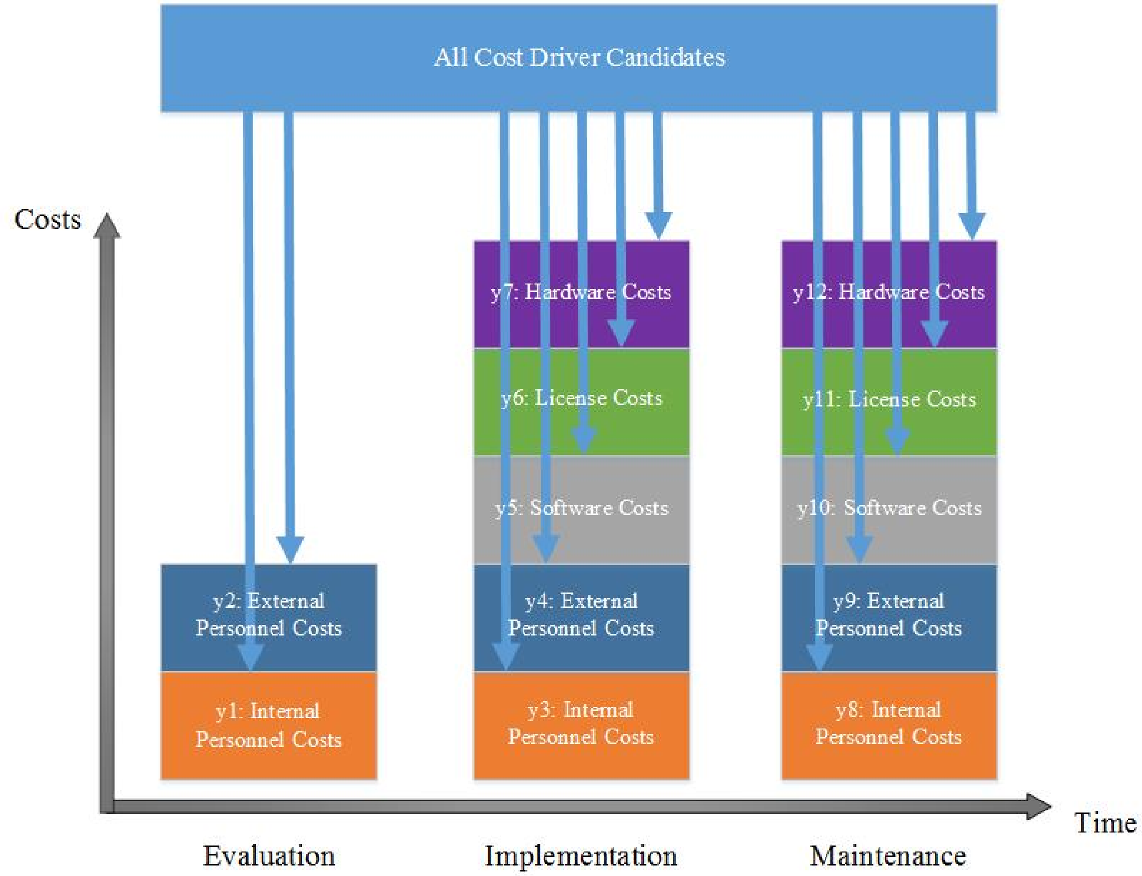
The timing component, which is the first dimension, is represented on the x-axis as the different lifecycle phases, which are the evaluation, implementation, and ten-year maintenance phase, which altogether reflect the complete ERP lifecycle.

The second dimension, which is the cost component, is represented on the y-axis. The costs for the three lifecycle phases are thus represented as three main bars in the graph. Every lifecycle bar consists of the five different cost types, which are internal personnel costs, external personnel costs, software costs, license costs, and hardware costs. For the evaluation phase, the cost types software costs, license costs, and hardware costs are excluded because they are not relevant during this lifecycle phase.

The combination of the lifecycle phases and the cost types results in a cost structure with 12 cost fields for an ERP lifecycle.

The cost drivers are the third dimension of the conceptual framework. The cost drivers represent the influencing factors that determine the extent of the twelve cost fields within the ERP lifecycle.

Figure 4.2: Conceptual framework derived from the literature review



This analysis should provide a first comprehensive and extensive overview of all emerging ERP costs. By establishing relationships between cost fields and cost drivers, ERP managers are equipped with a guideline for controlling their projects because they are aware of the factors that could influence the respective twelve cost fields.

On the basis of these findings, two approaches of assessing the cost estimation are tested. The first approach is to predict the costs for every single cost field. The second approach is to predict the total costs of the whole ERP lifespan and to distribute these costs according to the cost structure identified by answering RQ1.
Based on this conceptual framework, this thesis has addressed its research questions. They should access the research issue. The questions are presented in the next section.

4.6 RESEARCH QUESTIONS

The aim of this research should be met by answering research questions.
It was decided to break some questions down into smaller research sub-questions so as to receive more precise answers.
They are presented in the following three tables:

Table 4.5: Research question I

RQ1: What are the costs of ERP systems during their lifecycle phases?	**RQ1-1**: What are the costs of internal personnel (y1) and external personnel (y2) during the evaluation lifecycle phase? **RQ1-2**: What are the costs of internal personnel (y3), external personnel (y4), ERP software (y5), licence (y6) and hardware (y7) during the implementation lifecycle phase? **RQ1-3**: What are the costs of internal personnel (y8), external personnel (y9), ERP software (y10), licence (y11) and hardware (y12) during the ERP lifecycle maintenance phase? **RQ 1-4:** What are the costs of ERP systems during their whole lifespan?

Table 4.6: Research question II

RQ2: Which cost drivers influence ERP costs?	**RQ 2-1:** Which cost drivers influence internal personnel costs (y1) and external personnel costs (y2) in the evaluation lifecycle phase? **RQ2-2**: Which cost drivers influence internal personnel costs (y3), external personnel costs (y4), ERP software costs (y5), licence costs (y6) and hardware costs (y7) in the implementation lifecycle phase? **RQ2-3:** Which cost drivers influence internal personnel costs (y8), external personnel costs (y9), ERP software costs (y10), licence costs (y11) and hardware costs (y12) in the maintenance lifecycle phase?

Table 4.7: Research question III

RQ3:
How can the identified cost drivers and cost fields be used to predict ERP costs?

In order to answer the research questions, an empirical quantitative research design is selected which is described in chapter 5.

5. RESEARCH METHODOLOGY & DESIGN

This chapter explains how the research questions will be assessed and justifies the guidelines and methods applied in this thesis.
Looking at the research questions, the nature of this research is explanatory and predictive, aiming to both test relationships between variables and predict general relationships between costs or cost types and cost drivers.

With its deductive and deterministic nature of research, the postpositivist paradigm underlines this research.
Since paradigms often guide the applied methodology, this chapter explains and compares the two contrasting paradigms postpositivism and constructivism and gives a justification for the selected approach.

Having justified applying a purely quantitative methodology, section 5.2 starts to explain the nature of the methodology and gives an overview of the corresponding strategies of inquiry, including survey design, the scales and measurement of the relevant variables in the questionnaire, the sampling technique, and the data collection process.

In the next step, the quality criteria objectivity, reliability and validity will be discussed. Since this thesis involves sensitive data of people and organisations, the thesis needs to be aware of ethical considerations. This issue is reported in section 5.4. Finally, the chapter is summarised.

5.1 JUSTIFICATION OF RESEARCH PARADIGM

The selection of an appropriate "paradigm" or "philosophical world view" is an important step in undertaking a scientific research, because they give basic notions of the research design used later (Creswell, 2009; Easterby-Smith, Thorpe, & Jackson, 2008). The term "paradigm" or "world view" means a basic belief system or set of assumptions which cannot be proven or disproved but helps to provide a conceptual guideline to the researcher (Guba, 1990; Lincoln & Guba, 1985). All existing paradigms can be characterised by inquiring the way they refer to their ontological, epistemological and methodological assumptions (Guba, 1990).

Ontology is defined as "philosophical assumptions about the nature of reality" (Easterby-Smith, et al., 2008, p. 60) "[…] that raises questions of the assumptions researchers have about the way the world operates and the commitment held to particular views" (Saunders, Lewis, & Thornhill, 2007, p. 108). The underlying beliefs lead to different guidelines and principles regarding the question how research should be conducted, which strategies provide valid results in order to cover reality, and which methods are appropriate for generating knowledge.
The main contrasting positions here are the objectivistic (realistic) and subjectivistic views (Saunders, et al., 2007). The objectivist believes in one real world functioning consequently to natural rules in which social entities exist independently from the individuals living in it (Easterby-Smith, et al., 2008; Guba, 1990; Saunders, et al., 2007). By contrast, subjectivists assume that reality is created from the "perceptions and consequent actions of those social actors concerned with their existence" (Saunders, et al., 2007, p. 108).
The opposing positions lead to different assumptions as to what constitutes acceptable knowledge and how reality can be portrayed. This scientific level, succinctly "how reality can be known", is recognised as epistemology.

The term epistemology is defined as a "general set of assumptions about the best ways of inquiring into the nature of the world" (Easterby-Smith, et al., 2008, p. 60), which means to question "[…] what is the nature of the relationship between the knower (the inquirer) and the known (or knowable)" (Guba, 1990, p. 18). Which method provides the most appropriate approach to reality depends on the underlying ontology, and normally corresponds to it. Since the objectivist believes in one concrete external reality, this position assumes that reality can be measured by using objective methods that can reliably cover the world and its properties. On the other hand, the subjective worldview assumes that reality is created by individuals and therefore needs to be accessed differently. Since this ontology focuses on the different constructed realities, reality needs to be interpreted dependent upon the interpretations of the social actors. In contrast to objective epistemology, the interpretative approach does not believe in an independent observer who is able to measure reality, but assumes that the observer/inquirer is part of reality and needs to adequately reflect his/her position in the processes.

The underlying ontology and epistemology lead to corresponding methodologies and methods. Methodology means the "combination of techniques used to enquire into a specific situation" (Easterby-Smith, et al., 2008, p. 60). Since the objectivistic and interpretative worldviews have completely different epistemologies, it seems quite natural that they need to focus on different methodologies in order to inquire into the reality. Generally, objectivists use survey and experimental research to access reality, whilst interpretative researchers apply methodologies like grounded theory, action research and ethnography.

This discussion has shown that research and its results depend on the individual ontology, epistemology and methodology.

Many different paradigms are discussed in literature (Easterby-Smith, et al., 2008). When comparing different literature on "research philosophy and design", most contents focus on different paradigms. Only the contrasting paradigms positivism or rather postpositivism and constructivism could be found in nearly all of them.

This study summarises these two key paradigms and demonstrates their contrasts in table 5.1 before explaining them in more detail in the next two subchapters. A justification of the chosen paradigm is provided at the end of this section.

Table 5.1: Key paradigms basic beliefs

Paradigm	Ontology	Epistemology	Methodology
Postpositivism	Critical Realist: Reality exists and is ruled by natural laws - but reality cannot truly be perceived. Knowledge of the rules allows a time- and context-free generalisation.	Modified Objectivist: It is necessary to have a neutral position. Objectivity is not absolutely possible. Results are shaped by interaction between inquirers and inquired. This problem is redressed by using the critical theory and community. Findings are probably true.	Modified Experimental: Critical multiplism (triangulation) Falsification of hypothesis
Constructivism	Relativist: Reality is individually constructed and dependent on mental framework.	Subjectivist: Findings are the creation of interaction between inquirer and inquired.	Hermeneutic/ Dialectic: Depicting individual construction accurately and comparing or contrasting these existing constructions. The aim is to reconstruct one or more constructions.

Source: Adapted from Guba (1990, p. 20 et seqq.)

5.1.1 POSTPOSITIVISM

Postpositivists believe that the external existing reality cannot be accurately perceived by any existing research methods. Explanations of social relations, circumstances and phenomena are therefore imperfect and knowledge is consequently tentative (Creswell, 2009; Guba, 1990).

Postpositivist research methodology is defined as an empirical inquiry or traditional scientific approach (Creswell, 2009; Guba, 1990). That means “the principals and procedures that govern investigations of the physical world” (Schwandt, 1990, p. 259) are adapted to social disciplines in the notion that “[…] epistemologically, the two sciences are rather similar – the relationship between the evidence that is appealed to, and the knowledge claims that are made, is the same” (Phillips, 1987, p. 395).

Therefore their research design is primarily a deterministic, quantitative approach that aims to "determine effects and outcomes" and describe causal cause-and-effect relationships in order to seek generalisation of the findings (Creswell, 2009, p. 7; Guba, 1990; Lincoln & Guba, 1985). To determine these, postpositivists reduce their research interests into small sets, transform them into variables, formulate hypotheses and/or research questions and test them mainly with experimental or non-experimental designs for supporting or rejecting the hypotheses (Creswell, 2009; Easterby-Smith, 2008; Guba, 1990, Schwandt 1990). This procedure is defined as a reductionist and deductive approach (Burns & Burns, 2008; Creswell, 2009).

Aiming to balance the imperfection of any research method, postpositivists may use forms of triangulation to receive reliable data (Guba, 1990) by gathering data from many different sources or with different methods in order to indicate or check the findings. Furthermore, a preferably objective position is essential but not realisable in the postpositivistic paradigm.

Although the traditional methodology of the postpositivistic paradigm is rather a quantitative one, this strict approach is controversially discussed within the literature and reached its peak in the "Paradigm Wars" discussion (Tashakkorie & Teddlie, 1998). Whilst some scientists argued that mixing quantitative and qualitative methods and therewith postpositivistic and constructivistic paradigms is irreconcilable and untenable, others claimed the mono methods obsolescent and support the use of both quantitative and qualitative approaches (Creswell, 2009; Lincoln & Guba, 1985). However, nowadays it is possible to undertake qualitative designs within the postpositivistic paradigm (Guba, 1990).

5.1.2 CONSTRUCTIVISM

The constructivist paradigm starts from a relativistic ontology. Constructivists do not believe in any external reality, but assume that "individuals develop subjective meanings of their experiences – meanings directed toward certain objects or things" (Creswell, 2009, p. 8), which means for Lincoln (1990) that reality is a "set of holistic and meaning-bound constructions that are both intra- and interpersonally conflictual and dialectic in nature" (Lincoln, 1990, p. 77).

The aim of constructive research is therefore the identification of the multiple constructed realities, bringing them into consensus which is uttered in theories (Creswell, 2009; Guba, 1990), "webs of mutual and plausible influence expressed as working hypotheses, or temporary, time- and place-bound knowledge" (Lincoln, 1990, p. 77). The fact that theory arises from data in this paradigm can be classified as an inductive approach (Burns & Burns, 2008; Creswell, 2009; Saunders, et al., 2007).

The constructivist approach to capture the participants' world views is rather a qualitative, holistic, hermeneutic and dialectic one (Lincoln, 1990). To access them, research focuses on "social processes of construction, reconstruction and elaboration […]" (Lincoln, 1990, p. 78) in natural contexts by using methods like field research, case studies, ethnographic approaches, anthropological research, narrative methods and elements of grounded theory (Creswell, 2009; Schwandt, 1990).

Being aware that inquiry cannot be value-free and interaction between inquirer and inquired has an influence on the results, constructivists see this subjectivist epistemology as the only way to access the constructed realities of individuals (Creswell, 2009; Guba, 1990).

5.1.3 CONCLUSION

The nature of this research is mainly explanatory and predictive, aiming to both test relationships between variables and predict general relationships between costs or cost fields and cost drivers. With its deductive and deterministic nature of research, the postpositivist paradigm seems to be more appropriate, and will therefore be followed in this research. In detail, according to Creswell (2009) and Easterby-Smith et al. (2008), this means:

- The research idea will be reduced to a set of variables, controlled through design or statistical analysis.
- Relationships between variables will be tested through surveys or experiments.
- Although gathering objective data is not absolutely realisable, validity and reliability of scores and instruments lead to reliable findings.
- The aim will be generalisation.

- Quantitative research approaches will be used.

Elements of triangulation are often required when conducting a postpositivist research. Scientists such as Easterby-Smith et al. instead argue that they have “reservations about mixing methods when they represent very distinct ontologies” and […] “there are difficulties when different kinds of data say contradictory things about the same phenomena” (2008, p. 71). Sharing this opinion, this research will rely on a purely quantitative methodology.

5.2 RESEARCH DESIGN

The chosen paradigm leads mostly to a correspondingly applied methodology. This research follows the postpositivist paradigm and applies exclusively quantitative approaches.

A quantitative approach represents the traditional form of research to establish general laws and principles aiming to look for relationships between variables (Burns & Burns, 2008; Creswell, 2009; Gliner, Morgan, & Leech, 2009) and tries to “provide evidence that a so-called independent variable is caused by an observable change or difference in the dependent variable” (Gliner, et al., 2009, p. 45). To analyse relationships and evidences, researchers start by formulating a research question and/or hypothesis to shape and specifically focus the purpose of the research. This is already done and presented in chapter 4.

Compared with other designs, quantitative data is relatively objective and comparatively easy to classify and quantify (Gliner, et al., 2009). Table 5.2 highlights and contrasts the quantitative methodology to the qualitative and mixed-method approach, and their consequences for inquiry, methods, and practises of research.

Table 5.2: Quantitative, qualitative and mixed-method approaches

Tend to or Typically	Quantitative Approach	Qualitative Approach	Mixed-Method Approach
Paradigm	Postpositivist	Constructivist	Pragmatist
Employ these strategies of inquiry	Surveys and experiments	Phenomenology, grounded theory, ethnography, case study, and narrative approaches	Sequential, concurrent, and transformative
Employ these methods	Closed-ended questions, predetermined approaches, numeric data	Open-ended questions, emerging approaches, text or image data	Open- and closed-ended questions, emerging and predetermined approaches, and quantitative and qualitative data and analysis
Use these practices of research as the researcher	-Tests or verifies theories or explanations -Identifies variables to study -Relates variables in questions or hypotheses -Uses standards of validity and reliability -Observes and measures information numerically -Uses unbiased approaches -Employs statistical procedures	-Positions him- or herself -Collects participant meanings -Focuses on a single concept or phenomenon -Brings personal values into the study -Studies the context or setting of participants -Validates the accuracy of findings -Makes interpretations of data -Creates an agenda for change or reform -Collaborates with the participants	-Collects both quantitative and qualitative data -Develops a rationale for mixing -Integrates the data at different stages of inquiry -Presents visual pictures of the procedures in the study -Employs the practices of both qualitative and quantitative research

Source: Creswell, 2009, p. 17

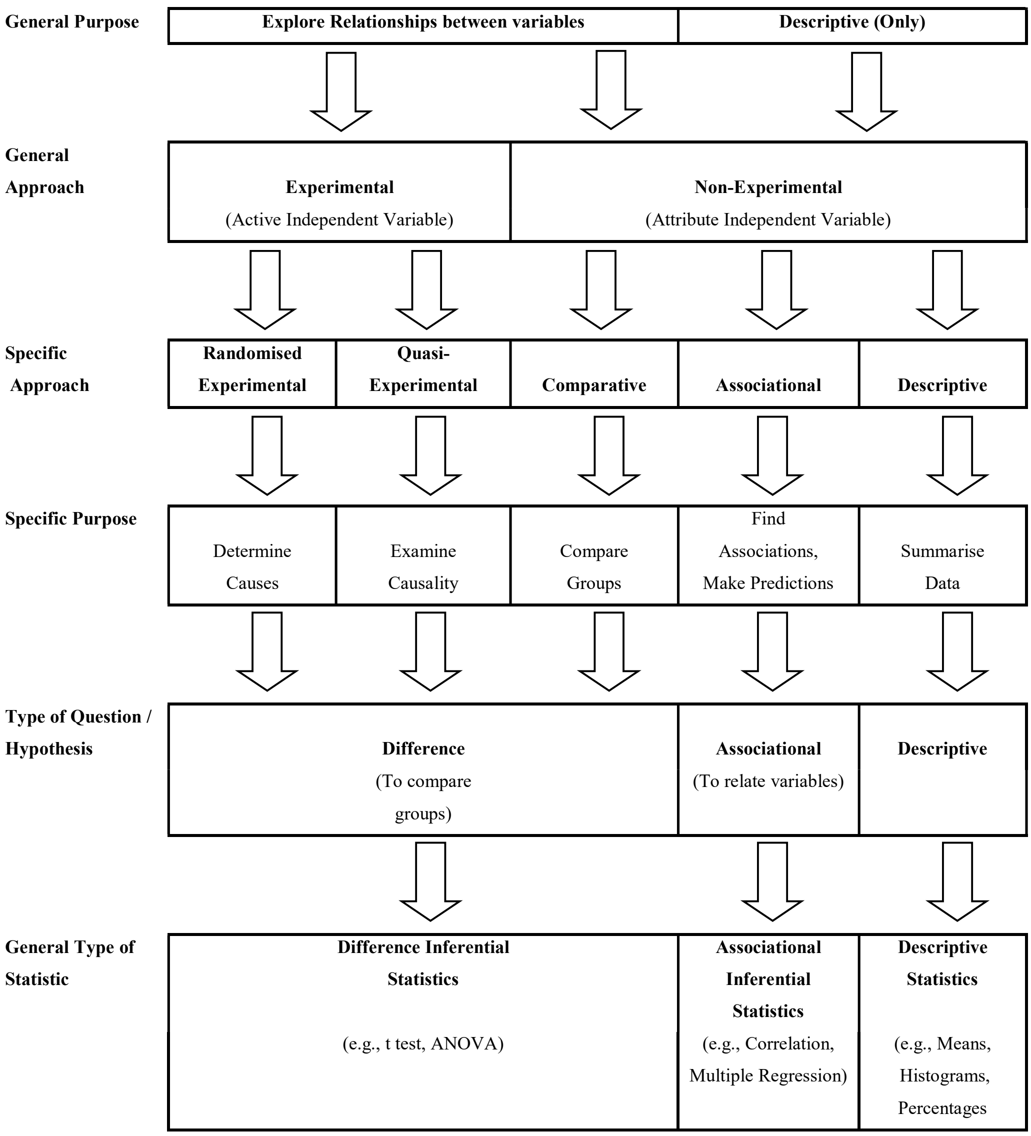
However, in order to ensure a successful answering of the research questions, the development of appropriate research design is necessary which includes several aspects which will be discussed in the following sub-sections.

5.2.1 STRATEGIES OF INQUIRY

Having selected a quantitative approach, the review will now consider the different approaches within this methodology before giving a justification for the chosen approaches.

Gliner et al. (2009) present a table that divides the two purposes existing in quantitative approaches into the two general experimental and non-experimental approaches within this research methodology. Figure 5.1 shows their scheme.

Figure 5.1: Quantitative approaches



Source : Gliner et al. 2009, p. 46

Since the nature and environmental conditions of ERP effort prediction are too complex to reproduce in experimental designs, a non-experimental approach is chosen for all of the presented research questions.

Non-experimental methods are categorised into comparative, associational and descriptive approaches and are based on surveys and observations.
The difference between the comparative and associational approaches is that comparative designs usually have a few categories of the independent variable and make comparisons between groups (Gliner, et al., 2009). In contrast, the independent variable in associational research is treated as continuous as if all "participants" of a study were in a single group (Gliner, et al., 2009). The differences and similarities are summarised in table 5.3.

Table 5.3: Comparison of the non-experimental research approaches

Criteria	Comparative	Associational	Descriptive
Random assignment of participants to groups by investigator	No	No (only one group)	No groups
Independent variable is active	No (attribute)	No (attribute)	No independent variable
Number of levels of independent variable	Usually 2-4	Usually 5 or more ordered levels	No independent variable
Relationships between variables or comparison of groups	Yes (comparison)	Yes (relationship)	No

Source: Gliner, et al., 2009, p. 47

RQ1 can be answered by descriptive statistics yielding the percentage distribution of the 12 cost fields. Seeking to find and measure relationships between cost drivers and cost fields and forecast the ERP costs, RQ2 and RQ3 have an explanatory and predictive nature of research. The associational inquiry is the only specific approach which is appropriate for answering them:

For the answering of RQ2, correlation analysis is chosen as the appropriate method to measure the relationship between the cost drivers and costs incurred during the single lifecycle phases. Correlation analysis is the standard method for the measurement of the degree of correspondence between variables (Burns & Burns, 2008).

For the answering of RQ3, multiple regression analysis will be used for the development and assessment of an ERP cost prediction model. Multiple regression is a method for determining the dependent variable from the value parameters of two or more independent variables (Burns & Burns, 2008).
Excel will be used to analyse the descriptive statistics and analysis of the cost structure in RQ1.
SPSS will be used for conducting the correlation analysis of RQ2. SPSS will also be used for conducting the regression analysis within RQ3.

After the different strategies of inquiry and their applied methods have been introduced and selected, the survey design will be discussed in the next section.

5.2.2 SURVEY DESIGN

Adopting the associational approach in this research, data is gathered through surveys (Gliner, et al., 2009). Surveys are defined as "a process of collection information from a sample of people who have been selected to represent a defined population" and are commonly used in explanatory and descriptive studies (Burns & Burns, 2008, p. 486). The aim of explanatory surveys is to examine relationships of variables without experimental manipulation and the generation of models containing those relationships (Burns & Burns, 2008; Saunders, et al., 2007).

Surveys can be divided into five different techniques: group self-completion, mailed self-completion, personal interview, telephone interviews, and internet-based surveys (Burns & Burns, 2008).

In group self-completion, a questionnaire is transmitted to a "group of people who are assembled in a certain place for a specific purpose such as a class […]" (Gliner, et al., 2009, p. 184). This technique is mainly used when "large samples can be gathered together" (Burns & Burns, 2008, p. 489). Not having those contexts, this method is not appropriate for this research.
Mailed self-completion is a method in which a questionnaire is mailed to a defined sample and the sample is requested to respond (Gliner, et al., 2009). Although this method has

many advantages, mailed questionnaires have the poorest response rate of all survey methods (Burns & Burns, 2008; Gliner, et al., 2009). Aiming to have the highest possible return rate, this method was not chosen for this research.
In personal (face-to-face) interviews, the researcher surveys the individuals of a defined sample directly (Burns & Burns, 2008; Gliner, et al., 2009). This method is relatively expensive and time-consuming (Burns & Burns, 2008; Gliner, et al., 2009). Therefore, this research design will actually not rely on this method.

In telephone interviews the inquirer interviews the individuals of a sample via phone. In comparison to face-to-face interviews, telephone interviews are inexpensive and faster to complete, but gathering sensitive data with this method is questionable (Burns & Burns, 2008). Furthermore, they have to be brief and simple (Burns & Burns, 2008; Gliner, et al., 2009). Since the focus is on sensitive data, which might not be readily available to the inquired, the telephone does not seem to be an appropriate method for this research.

Internet-based surveys are becoming more and more popular and allow the questionnaire to be set up with special online programs (Burns & Burns, 2008; Gliner, et al., 2009). The advantages and disadvantages are shown in table 5.4.

Table 5.4: Advantages and disadvantages of internet surveys

Advantages	Disadvantages
Continuous collection of data as long as you want to run it	Potential for multiple submissions from individuals
Specific sample from lists	General invitations can lead to biased samples
Economical	Technical issues can lead to loss of data, and respondent frustration, e.g. network speed
Highly motivated voluntary participants	Little control of experimental settings or sampling
Increased generalisability of findings	Considerable time and cost expenditure for designing and developing the survey, its layout, and maintaining the website
Decreased drop-out	Ethical issues, such as informed consent, withdrawal buttons, and debriefing
Expense and time saved on entering data as already electronically recorded	Data storage issues
Cross-cultural research facilitated	

Source: Burns & Burns, 2008, p. 497

When comparing the advantages and disadvantages of this method, the author feels that the advantages dominate with regard to this research.
Especially the fact that targeting specific professionals is possible through an email mailing and considering economical components with respect to time and cost, lead to the decision to collect data via an online survey.

5.2.3 VARIABLES

In order to collect, process, and analyse data, it is important to define a detailed research design which is geared to support answering the research questions. This includes the definition of variables and their types of measurement.

Both ERP cost fields and cost drivers were introduced in section 4. The cost drivers are the independent variables and called x variables (x_i). The cost fields are the dependent variables and labelled as y variables (y_i).

A detailed description of the independent as well as the dependent variables and their levels of measurement are given in the following.

INDEPENDENT VARIABLES

35 different cost drivers categorised in three dimensions have been introduced for the theoretical framework. Their measurability needs to be distinguished into latent and manifest variables. While manifest variables can be directly observed or measured, latent variables cannot be directly observed or measured. Table 5.5 presents an overview on the next page.

The analysis identified 13 independent variables as manifest, which can be measured in ratio scales[15]. They can be easily inquired and simply quantified by their nature (e.g. Variable x2: How many organisational departments exist within your company).

[15] Variable scales are described in Gliner et al. (2009) and Kuehnel & Krebs (2004). The variables used in this research are classified according to them.

Table 5.5: Measurability of cost driver variables

Organisational Cost Drivers		**Technical Cost Drivers**		**Situational Cost Drivers**	
Manifest	**Latent**	**Manifest**	**Latent**	**Manifest**	**Latent**
x_1: No. of locations		x_5: ERP system	x_{13}: Complexity of data	x_{16}: No. of int. project members	x_{19}: Fit of system/ organisation
x_s: No. of organis. units or depts.		x_6: No. of interfaces	x_{14}: Complexity of interfaces	x_{17}: No. of external consultants	x_{20}: Team quality
x_3: No. of total employees		x_7: No. of modifications	x_{15}: Complexity of reports	x_{18}: Ratio external/ internal	x_{21}: Team maturity
x_4: Revenue		x_8: No. of reports			x_{22}: Team composition
		x_9: No. of EDIs			x_{23}: Availability of management
		x_{10}: No. of total users			x_{24}: Availability of business users
		x_{11}: No. of user groups			x_{25}: Consultant experience
		x_{12}: Modules			x_{26}: Consultant quality
					x_{27}: Critical attitude of users
					x_{28}: User quality
					x_{29}: Employee involvement
					x_{30}: Management involvement
					x_{31}: Maturity of processes
					x_{32}: Complexity of processes
					x_{33}: Stability of organisation
					x_{34}: Willingness to change
					x_{35}: Commitment management

In contrast, latent variables cannot be quantified that easily. Often they need to be measured by constructs of manifest variables or are determined by value parameter when asking the inquired for an opinion. This kind of variable is mostly called opinion variable in the literature (Dillman, 2000).

Opinions are mostly measured by applying the so-called Likert scale which rates valuations in pre-defined scales, e.g. from - 5 to 5 (Burns & Burns, 2008; Gliner et al., 2009; Diekmann, 2004).

The survey was addressed to ERP experts who were able to make correct valuations and reliable statements. This is why, in this survey design, it was decided to access the latent variables by asking the inquired for their opinions instead of constructing manifest variables.

Having chosen a closed-ended design, these variables need to be transformed into rating scales. In this research, the statement of attitude should be estimated in a scale from 1 to 10.

Category 1 means a strong disagreement with the statement or the most negative estimation of a variable category. 10 means a strong agreement or the most positive estimation of a variable category (e.g. Variable x25: Please estimate the experience of the consultants: 1 (not experienced) […] to 10 (very experienced).

A 10-point Likert scale was chosen for several reasons.

Firstly, participants should be forced to not position neutrally by applying an equal point Likert scale. Since the neutral position 5.5 is not selectable, a participant must decide on 5 or 6 and thus between a more approving and a more disapproving position.

Another reason for the selection of a 10-point scale is that 10 points allow a certain differentiation of reply possibilities, which might increase the accuracy without overdoing it and thereby negatively influencing the handling.

A further reason is that people are used to ten-point scales due to the decimal system. Since the Likert scale in this study is regarded as an interval scale, a tenner scale supports the cognition of equidistant steps, meaning that the space between the points is equal, which is really important for generating a reliable analysis.

The scale of those “Likert” variables is controversially discussed in the literature. While some authors argue that the Likert scale does not have more than an ordinal scale (Burns

& Burns, 2008), others claim that in case of symmetric, numbered response options the Likert scale can be seen as quasi-quantifiable and therefore can be used as an interval variable (Blumberg, Cooper, & Schindler, 2005 cited in Saunders et al., 2007).
In this research, Likert variables will be seen as variables with interval scale. To justify this approach, the inquired will be advised that the value parameters are equidistant and the questionnaire will be visually supported by designing the questionnaire in a corresponding manner. Also, the choice of a 10-point scale supports the cognition of an equidistant scale.

DEPENDENT VARIABLES

The cost fields are the dependent y variables in this study. Table 5.6 repeats the overview given in section 5 and extends it by the category "measurement".

Table 5.6: Dependent variables

Dependent variable	Lifecycle Phase	Cost Type	Measurement
y_1	Evaluation	Internal personnel costs	Person-days or Euros
y_2	Evaluation	External personnel costs	Person-days or Euros
y_3	Implementation	Internal personnel costs	Person-days or Euros
y_4	Implementation	External personnel costs	Person-days or Euros
y_5	Implementation	ERP software costs	Euros
y_6	Implementation	Licence costs	Euros
y_7	Implementation	Hardware costs	Euros
y_8	Maintenance (10 years)	Internal personnel costs	Person-days or Euros
y_9	Maintenance (10 years)	External personnel costs	Person-days or Euros
y_{10}	Maintenance (10 years)	ERP software costs	Euros
y_{11}	Maintenance (10 years)	Licence costs	Euros
y_{12}	Maintenance (10 years)	Hardware costs	Euros

The variables "internal personnel costs" and "external personnel costs" are measured in person-days ERP or thousands of euros, depending on which format the participants are more comfortable with. To create a certain comparability, the person-days need to be converted to euros. This is presented in detail in chapter 7.1
The variables "software costs", "licence costs", and "hardware costs" are measured in thousands of Euros.

Thus all dependent variables are ratio scales.

5.2.4 DESIGN OF QUESTIONNAIRE

As discussed in section 5.2.2, this research collects data through an online survey providing self-completion questionnaires. Therefore, an appropriate online survey tool was required.

After comparing several different online survey tools, it was decided to use QuestBack's EFS Survey. EFS Survey enables researchers to easily design their own surveys using custom templates and providing a lot of graphical support for the questions. Furthermore, EFS Survey supports the whole invitation and reminder process. EFS also supports the conducting of anonymous surveys so that participants cannot be retraced. Results can then be downloaded into spreadsheet tools, such as Excel, or professional statistics programs, such as SPSS, for further analysis.

However, the presentation of internet-mediated questionnaires places higher demands on the design as opposed to that of postal questionnaires. In contrast to those questionnaires, internet-mediated questionnaires must consider the fact that respondents using different operating systems, browsers, and display screens may result in a different display of the questionnaire (Dillman, 2000 cited in Saunders, et al., 2007). Hence, it was important to use only simple components and a rather low solution of the questionnaire website to ensure that in all cases respondents would not have any presentation or navigation problems during the completion.

The questionnaire for this research consists of 49 questions and is presented in Appendix III (German original version). It shows that the questions are categorised into 8 different sections.
Besides the salutatory screen which explains the structure of the questionnaire at the beginning and a "thanking for the attendance" section at the end, the questionnaire consists of an additional six thematically structured sections:
Section A: Company details (organisational cost drivers)
Section B: Cost structure ERP evaluation

Section C: Cost structure ERP implementation

Section D: Cost structure ERP maintenance

Section E: Technical cost drivers

Section F: Situational cost drivers

For the sections A-D, text fields were used as the input method for every question. The input format was not restricted because a restriction could have led to the withdrawal of participants. Since the number of responses was limited, the gained data could be verified and harmonised after the closing of the survey.
As far as possible, the use of 'do-answer-checks' was avoided also because participants might interrupt the survey if they cannot go on without omitting one question. A 'do-answer-check' was only included for answers concerning costs being spent on evaluation, implementation, and maintenance, collected in sections B, C, and D, meaning that answers to these questions were mandatory. The decision to enforce entry of these variables was motivated by the fact that these variables are the dependent variables.

Sections B to D of the questionnaire contained questions regarding the dependent y variables. Since costs of internal and external personnel are often not known, the participants had the opportunity to enter these costs in person-days as an alternative.

Sections E and F asked for the value parameters of technical and situational cost drivers. With different types of questions, the section also had three different answer formats.
For example, the questionnaire provided text fields for questions like "How many users".

Another response format was the application of checkboxes for questions that required the participants to select the stated ERP modules implemented in their company. The option to choose from pre-defined answers enables the harmonisation of answers and reduces the risk that a participant forgets to mention a module. In case an implemented module was not mentioned in the pre-defined set of answers in the questionnaire, the respondents had the opportunity to enter additional modules in text fields.

The third response format was the 10-point Likert scale which was used to ask the opinion questions. The online survey tool supported the visual presentation of this question type in the form of a slide bar. This was complemented by the use and presentation of an opposing pair of values. In addition, every point of the 10-point scale was presented on the bar, which supports the cognition of an equidistant scale, meaning that the spaces between the points are equal as implied by an actual interval scale. This was also indicated in the FAQ section of the invitation email.

The original questionnaire is presented in Appendix III. Since the target group was made up of German companies, the questionnaire was formulated in German. A translation of the German questionnaire into English is presented subsequently (English Translation Appendix IV).

5.2.5 POPULATION AND SAMPLING

According to Gliner et al. “sampling is the process of selecting part of the larger group of participants with the intent of generalising from the sample (the smaller group) to the population (the larger group). To make valid inferences about the population, (it is necessary to) select the sample so that it is representative of the total population” (2009, p. 115).

This research aims to achieve that by applying the following steps suggested by Burns & Burns (2008) and Gliner et al. (2009):

I. Defining the target population
II. Identifying the sampling frame
III. Selecting and applying a sampling technique
IV. Defining the sample size

I. DEFINING THE TARGET POPULATION

This research focuses on German SMEs that have implemented an ERP system, employ 30 to 1,500 employees worldwide, and produce products in the industrial sector. The questionnaire has to be completed by an ERP specialist or executive staff employed in the company.

II. IDENTIFYING THE SAMPLING FRAME

According to Burns & Burns (2008), sampling frame means ‘a full list of the target population’. In this research, the population should have been accessed with the support of the German Chamber of Commerce providing a full list of contacts according to the above-mentioned characteristics of the target population. The procedure in which the researcher has ‘access to names in the population and can sample the population’ is called a single-stage technique (Creswell, 2009).

Unfortunately, the Chamber of Commerce was not able to provide such a list for reasons of data protection and no other organisation or authority collects such data in Germany. In consequence, this research needed to find an alternative way of generating an address pool of the target population, and used address files from different commercial and non-commercial providers in order to create a sample with the above-mentioned population criteria.

In total, this study managed to gather 1,900 address files of those organisations. This sample design has some weaknesses and limitations which will be discussed later.
Adding the 5 companies from the authors’ context who agreed to do the first pre-test in the survey design, the quantity of the full sample design is n = 1905 finally.

III. SELECTING AND APPLYING A SAMPLING TECHNIQUE

The sampling technique pre-defines the process and method of selecting the sample from the target population (Diekmann, 2004). Basically, the sampling techniques can be divided into two groups:

I. Probability or representative sampling designs, and
II. Non-Probability or judgemental sampling designs

(Easterby-Smith, et al., 2008; Saunders, et al., 2007).

Principally, probability samples are designs in which the inclusion probability of participants or elements is known. The participants typically have an equal and independent chance of being included in the sample (Easterby-Smith, et al., 2008; Gliner, et al., 2009). Since the relationship between the sample and the target population is precisely known,

probability designs are the most representative techniques and are therefore ideal for survey designs and inferences (Black, 1999; Saunders et al., 2007).
Probability sampling comprises different methods, namely simple random sampling, systematic sampling, stratified sampling, cluster and multi-stage sampling (Kalton, 1983).

Non-probability designs, in contrast, do not know the inclusion chance of participants or elements and therefore provide some difficulties with generalisation of statistical findings (Easterby-Smith et al., 2008; Saunders et al., 2007). These designs have their origins in several "practical problems that researchers have encountered in carrying out their work" (Easterby-Smith et al., 2008, p. 218).
Common non-probability sampling techniques are convenience, self-selection, quota, purposive and snowball sampling (Easterby-Smith et al., 2008; Saunders et al., 2007).

Since this study is confronted with the practical problem of generating a full list of the target population, it cannot apply the desired simple random probability method.

Instead it needs to follow one of the non-probability techniques and chose the self-selection sampling method. It will be explained in the following.

According to Saunders et al. (2007), the self-selection sample involves two steps:

I. Communicate the need for participants
II. Collect the data from those who respond and check their relevance.

In the first step, the potential respondents need to know about the study. In this study, the 1,900 address files gathered from commercial and non-commercial sources will be used. Letting them know involves some kind of advertising or promotion. Promotion can have different forms and might include articles and advertisements in print media, on the radio, an online notice board, or some other medium (Lund Research, 2012; Saunders et al., 2012).
Since not all respondents might be relevant to the research or do not match the defined population, the applicants need to be checked and either included in the sample or sorted out in a second step (Lund Research, 2012).

In this study, invitation letters which promoted the participation in this research were sent to all addresses gathered from commercial and non-commercial organisations.
The original invitation email is presented in Appendix V (original German version). Since the target group was German companies, the questionnaire was formulated in German. A translation is presented subsequently (English Translation in Appendix VI)

The presented selection of the target group and sample technique is subject to some restrictions. The main weakness of this study is the sample technique. Since a self-selecting sample was used, the respondents might not be representative. To ensure this, a random sampling method should have been applied instead. Unfortunately, this was not possible within the research. The addresses of potential respondents were obtained from different commercial address services, which cannot ensure that the provided contacts will portray an accurate sample of the sample frame. Nevertheless, it was possible to select the obtained addresses according to the criteria that the companies (a) are German, (b) come from the industrial sector, and (c) have 30 to 1,500 employees. Consequently, this research might have limitations in terms of external validity and generalisability.
Furthermore, the findings of the study are restricted to German SMEs with 30 to 1,500 employees having their core business the industrial sector. It is questionable whether the results of this thesis can be transferred to other business sectors, other countries, and companies with different employee cultures and structures.
A further check was done after closing the response deadline to determine if the responses fit the target group, whether the respondents had already implemented an ERP system, and if they were German SMEs with 30 to 1,500 employees in the industrial sector.

IV. SAMPLE SIZE

The sample size and the number of responses certainly influence the quality and accuracy of statistical inferences. Thus, considering them is an important step in every research.

As mentioned above, this study is not able to apply the initial concept of random design. Instead of knowing the actual size of the target population and creating a sample out of it, this study managed to gather 1,905 potential respondents who are seen as the sample size in this research.

The literature discusses different approaches for determining the number of replies necessary to make valid inferences. Two of the most common approaches are the sampling error and the risk level which derives from the margin of error. However, those approaches are developed for random sampling methods (Israel, 1992) and will therefore not be followed.

Instead, this research needs to find another way of determining the number of respondents out of the sample size. The literature makes different suggestions to approach that.

Gliner (2009) and Saunders et al. (2007), for example, recommend a minimum of 30 valid participants for making accurate statistical inferences.

Another way of defining the necessary number of participants is to define a response rate. However, the minimum level for adequate response rates is controversially discussed in the literature and lacking consistency (Baruch & Holtom, 2008). While Baruch & Holtom (2008) present authors who suppose relatively high binary quantities of 50 to 80%, Sheehan (2001) found that the average response rate to email surveys is decreasing and reported a mean response rate of 24% in the year 2000. Other authors, like Cook, claim that the response representatives are much more important than the response rate (Baruch & Holtom, 2008).

However, for this study, even a response rate of 24% seems to be unachievable. To focus on ERP costs means to not only concentrate on very sensitive and intimate data, which are difficult to access, but also on the willingness to spend time looking for complex data, like former invoices, etc. The questionnaire for this research cannot be answered intuitively, but is bound to need some research on the part of the inquired himself/herself. Some organisations might even impose duties to maintain secrecy which would additionally hamper this field of research.

In order to obtain a realistic expectation of the response rate, an online pre-test, which is described in more detail in the next chapter, was conducted.
200 organisations were invited to participate in the survey, and 9 responded. Unfortunately, 2 of the 9 completed questionnaires were ineligible and thus invalid.

$$Total\ Response\ rate = \frac{Total\ Responses}{Sample\ size}$$

$$Valid\ Response\ Rate = \frac{Total\ Responses - Invalid\ Responses}{Sample\ Size} \rightarrow \frac{9-2}{200} \times 100 = 3.5\%$$

Assuming that a valid response rate of 3.5% is a realistic parameter, this study expects to generate 66 eligible responses out of the potential respondents. Since this expected number of respondents is higher than the recommended number of 30 participants, the result should be sufficient to make accurate statistical inferences in the following research.

The next section provides a detailed description of the data collection process and the instrumentation of the survey.

5.2.6 DATA COLLECTION

As discussed in section 5.2.2, this research will collect data through an online survey providing self-completion questionnaires. This chapter describes the data collection process.

Saunders et al. (2007, p. 386) strongly recommend conducting a pre-test prior to data collection in order to "refine the questionnaire so that respondents will have no problems in answering the questions and there will be no problem in recording the data". According to them, a pilot test enables a certain assessment of the data reliability and validity. Saunders et al. (2007) suggest starting with consulting experts about the structure and the content of the questionnaire. Their comments would allow making the necessary amendments before starting the pilot testing.

This study follows Saunders suggestion. The procedure is graphically described in table 5.7. It highlights the process of pre-testing the survey and the actual data collection process in three steps.

Table 5.7: Data collection process

Step	Step I Pre-Test	Step II Pilot Test	Step III Survey	Total
Quantity	5 Companies	200 Companies	1700 Companies	1905 Companies

Initially, five companies, which had agreed to participate in the survey, were consulted. They commented on the initial questionnaire and regarded the following aspects proposed by Bell (2005) (cited in Saunders et al., 2007, p. 387):

- how long the questionnaire took to complete
- the clarity of instructions
- which, if any, questions were unclear or ambiguous
- which, if any, questions the respondents felt uneasy about answering
- whether in their opinion there were any major topic omissions
- whether the layout was clear and attractive
- any other comments.

Based on the outcome of the respondents' comments, several amendments were implemented in terms of structure and ambiguities in order to optimise the questionnaire. Some questions in the study were re-defined based on the feedback of the participants. The answers to questions raised by the participants during the completion of the questionnaire were incorporated into the FAQ section of the invitation letter. In addition, a progress bar was added because the participants of the pre-test reported that they felt uncomfortable not being informed about the progress when answering the questionnaire, which could have lead to an abort before full completion.

After the questionnaire had been adjusted, the same five companies were pleased to fill in the questionnaire again, and once more after three weeks, in order to measure the reliability of the questionnaire. The reliability criterion in quantitative analysis is regarded in more detail in the next section and includes the results of the questionnaire reliability.

Afterwards, the second step of the pilot test could be heralded to start. Its aim was to test the data collection process and the response rate.

To this end, an appropriate online survey tool was required. It was decided to use the online survey tool EFS Survey in this research since there is not limitation as to the number of questions; 1,000 responses per month are permitted, and it is relatively inexpensive. Researchers using EFS Survey can easily design their own surveys using custom templates and post them on a website (Creswell, 2009). Results can then be downloaded into spreadsheet tools, such as Excel, for further analysis (Creswell, 2009).

200 companies were invited to participate in the pilot-test. These companies were randomly picked out of the sample and invited by mailing out the invitation mail.

According to Burns & Burns (2008), invitation mails should include the following aspects:

- Clear purpose of the research
- Clear communication why participants should help with this survey
- Define that ERP specialists or executive staff should complete the questionnaire
- Motivating section. It was pointed out that all valid respondents would receive a descriptive research report of the research findings and an individual benchmark for their companies' ERP implementations concerning the invested efforts.
- Introduction on how to participate in this survey (providing a hyperlink and access data)
- Pointing out how much time it will take to complete the survey.

Since questioning companies about their ERP expenditures is a very sensitive topic, the anonymous character of the survey seems to be important information. Therefore, this character was highlighted in the invitation email as well.
To avoid ambiguity and misunderstandings, the invitation email contained an additional detailed explanation of the procedure for answering the online questionnaire and an extensive FAQ section for potential upcoming questions.

This invitation email was addressed to the managing directors of the companies who were requested to participate in the survey or forward this email to an appropriate position within the company.

The respondents and their behaviour were analysed. This included aspects such as at which stages of the questionnaire the survey registered most interruptions, reasons making a response invalid, and the expectable response rate.

The results of the pilot-test showed that the data collection procedure delivers valid responses and, as mentioned in the previous chapter, produced seven valid responses, which means a response rate of 3.5%. The possibility of increasing the responses by sending a reminder was not examined in the pre-test. However, the analyses of the respondents showed a good understanding of the questions without finding serious difficulties. In consequence, the survey design could remain and the data collection process could get started.

The data collection process of the survey was carried out in a three-part mailing process consisting of a pre-survey contact (invitation email), a first reminder, and a second reminder. The invitation email was sent to the remaining 1,700 contacts of the sample.
Together with the invitation email, every participant received her/his own link to the online questionnaire which was valid until completion of the questionnaire. The link could not be matched to the individual companies because it was coded by the selected EFS Survey program. Despite the coding, the feature enabled a certain control of the avoidance of multiple responses by one respondent. The data were collected over a period of four weeks.

After this time had elapsed, reminder emails were sent to all participants. The issue with anonymous questionnaires is that they cannot be divided between respondents and non-respondents (Saunders, et al., 2007), therefore all participants received these reminder emails.
To account for this fact, the reminder email was formulated in such a way that early respondents were thanked and non-respondents were reminded to respond.
After a further two weeks, a second reminder email was sent. Two weeks after that, the data collection process was closed and no further responses were possible.

5.2.7 RESPONSE RATE

In this study, the sample size includes 1,905 potential respondents. After conducting the pilot test with 200 companies, a response rate of 3.5 % seemed to be a realistic expectation value. Table 5.8 gives an overview of the number of respondents in every step of the data collection process:

Table 5.8: Total response rate

Step	Number of Responses	Cumulative Number of Responses
(1) Pre-test 1	5	5
(2) Pilot Test	9	14
(3) Survey	66	80
(4) Reminder 1	4	84
(5) Reminder 2	1	85
Total	**85**	**85**

In total, this study managed to collect 85 responses during the data collection process. This corresponds to a total response rate of 4.5%.

As mentioned in the chapter Data Preparation, some responses were ineligible and had to be removed. In sum, 13 responses were excluded since they either did not match the defined target population or contained errors, like apparently unrealistic or implausible answers. Consequently, the number of respondents who were actually in the sample was 72. This corresponds to a valid response rate of 3.8%.

This value is close to the expectation value of 3.5% which resulted from the pilot test with 200 companies. In fact, it is exceeded by 8.5%. As mentioned earlier, the actual number of respondents should at least have a quantity of 30 valid ones in order to make accurate statistical inferences, according to Gliner (2009) and Saunders et al. (2007). With 72 valid responses, this criterion is matched.
However, a response rate of 4.5%, respectively a valid response rate of 3.8%, must be assessed as relatively low, which could have an effect on the survey results.

5.2.8 DESCRIPTIVE STATISTICS OF THE SAMPLE

With 72 responses, the characteristic attributes of this actual sample should be described briefly to show a more concrete picture of the respondents.
Since the target population was defined as German SMEs that have implemented an ERP system, employ 30 to 1,500 employees worldwide, and produce products in the industrial sector, this chapter will look into the distribution of the number of employees of the inquired companies, the sector in which they operate, and which ERP vendor they actually use.

Figure 5.2 shows the distribution of the company size based on the companies' total number of employees.

Figure 5.2: Companies by number of total employees

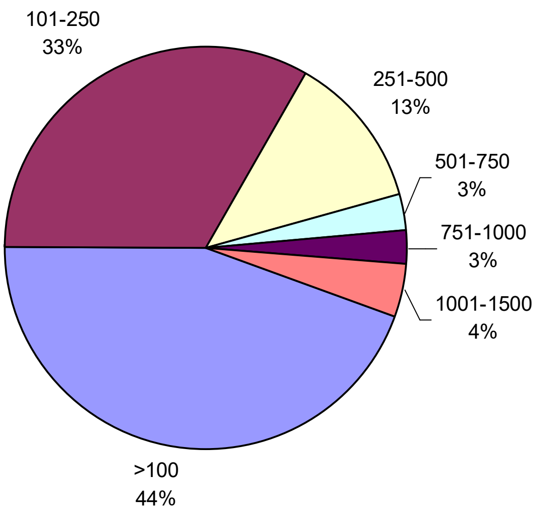


Figure 5.2 shows that the companies employing less than 100 employees make up the biggest portion with 44%. 33% of the sample employs 101 to 250 people, and 13% have between 251 and 500 employees. The remaining 10 % employ more than 500 people.

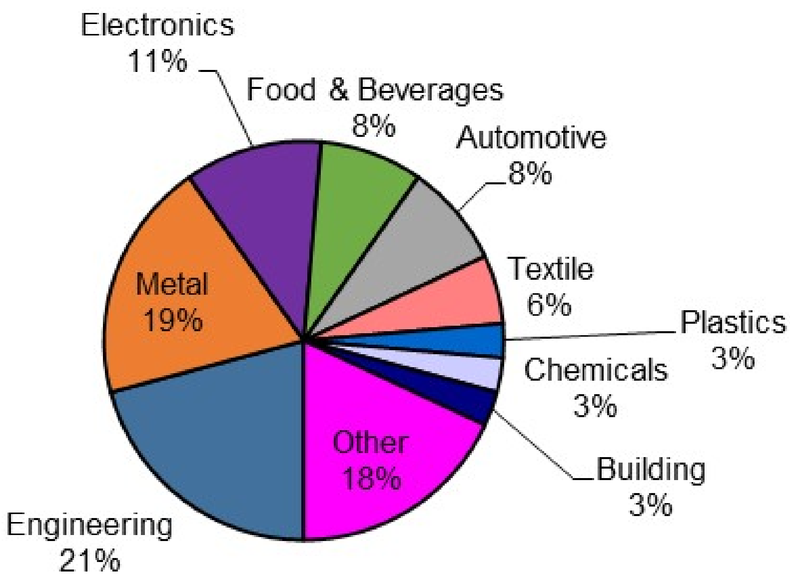
Table 5.9. presents the distribution of the company size also in absolute values.

Table 5.9: Companies by number of total employees in absolute values

Company Size by Number of Total Employees	Number of Responses	Cumulative Number of Responses
>100	32	32
101 – 250	24	56
251 – 500	9	65
501 – 750	2	67
751 – 1,000	2	69
1,001 – 1,500	3	72
Total	**72**	**72**

Aside from the number of employees, the industrial sector was another criterion for being in the target population. The industrial sector is very broad and can be differentiated into more sub-categories in order to map a more detailed picture. This is done in figure 5.3.

Figure 5.3: Distribution of respondents by industrial sector



With 21%, the largest number of inquired companies operates in the engineering sector. This sector is followed by the metal manufacturing sector with 19% and the electronics industry which, with 11%, is the third largest group within the sample.

These three sectors represent more than 50% of the sample.

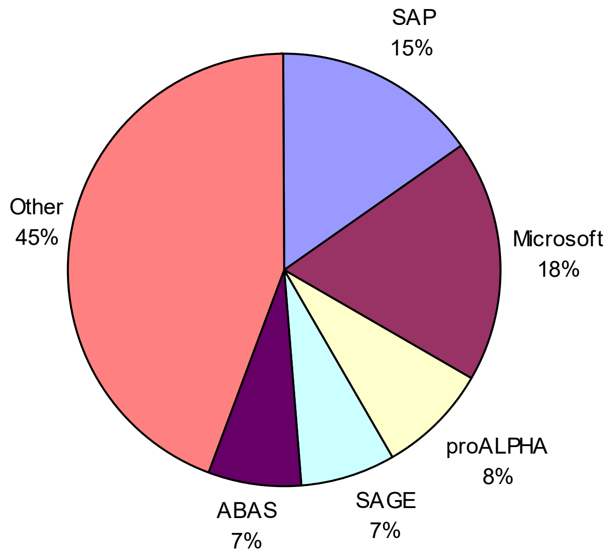
Table 5.10 presents the distribution of the respondents by industrial sector also in absolute values.

Table 5.10: Companies by industrial sector in absolute values

Industrial Sector	**Number of Responses**	**Cumulative Number of Responses**
Engineering	15	15
Metal	14	29
Electronics	8	37
Food & Beverage	6	43
Automotive	6	49
Textile	4	53
Plastics	2	55
Chemicals	2	57
Building	2	59
Other	13	72
Total	**72**	**72**

The third characteristic attribute to be examined is the ERP vendor. Data analysis showed that five ERP providers are the most dominant ones within this study. Figure 5.4 shows them in detail.

Figure 5.4: Deployed ERP systems of respondents



At 18%, Microsoft, with its ERP systems Dynamics AX and NAV, is the most frequently used ERP provider within this study. The second most common vendor is SAP, followed by proALPHA and, with the same values, ABBAS and SAGE.

ERP vendors who were named less than twice were summarised into the group "Other". With 45%, this category is far and away the most frequent one within this study.

Table 5.11 presents the ERP systems deployed by the respondents also in absolute values.

Table 5.11: Deployed ERP systems of respondents in absolute values

Deployed ERP Systems	Number of Responses	Cumulative Number of Responses
Microsoft	13	13
SAP	11	24
proALPHA	6	30
SAGE	5	35
ABAS	5	40
Other	32	72
Total	**72**	**72**

In all research, especially when measurements are obtained by using instruments like self-administered questionnaires, it is essential to know what faith and conclusions can be put into the data received (Burns & Burns, 2008). Therefore, the concepts of objectivity, reliability and validity, also known as quality criteria, will be discussed in the following three sections.

5.3 QUALITY CRITERIA

Objectivity, reliability, and validity are the most important quality criteria for measurements within research. The following three sub-sections report how these quality criteria are applied in this study and point out possible limitations and weaknesses of the data quality.

5.3.1 OBJECTIVITY

Degree of objectivity means to what extent a research outcome is independent of any kind of influences outlying to the subject of investigation. This denotes that objective research outcomes must be independent of the person conducting the research. In other words, objectivity is given when different persons come to the same results by using the same measuring instrument (Diekmann, 2004). According to Lienert & Raatz (1998), objectivity can be differentiated into three sub-categories of objectivity, which are:

- Implementation objectivity
- Evaluation objectivity
- Interpretation objectivity.

IMPLEMENTATION OBJECTIVITY

Implementation objectivity refers to the constancy of research conditions. It can be affected when research conditions are subject to disturbing factors. In this context, this means that implementation objectivity will be harmed if not all respondents are surveyed under comparable research conditions (Rammstedt, 2004). According to Rammstedt (2004), implementation objectivity can be affected by:

- Interview effects
- Item order effects
- Liability to present individual atmosphere
- Discontinuation during answering of the questionnaire

Implementation objectivity can be ensured through the use of maximum standardised data ascertainment (Rammstedt, 2004).

Due to the absence of an interviewer with self-completion online questionnaires applied in this research, and the highly standardised questionnaire and data collection process, very high implementation objectivity is provided in this research.
Theoretically, interferences with implementation objectivity in internet-mediated questionnaires can occur due to the application of different operating systems, browsers, and display screens of the participants, which may result in different displaying of the questionnaire (Dillman, 2000, cited in Saunders et al., 2007).
According to Welker, Werner, & Scholz (2005), implementation objectivity therefore decreases with increasing technical complexity of an online questionnaire. Therefore, only simple graphical components and a rather low resolution of the questionnaire website were applied. This ensured that respondents would not have any presentation or navigation problems during the completion and that the questionnaire was displayed similarly for all respondents.

EVALUATION OBJECTIVITY
Evaluation objectivity designates the degree of independence between data evaluation and its evaluator. In particular, such issues can occur during the encoding of answers to open questions, the conversion of a respondent’s given verbal answer into ticking off an answer on the questionnaire, or the errors during electronic data entry of questionnaires previously recorded in writing (Rammstedt, 2004). Evaluation objectivity means that results from the data evaluation of one evaluator must be equal with the results of another evaluator.

Evaluation objectivity was ensured in several ways in this research.
Firstly, through application of the standardised data ascertainment and analysis software, it was ensured that collected data were evaluated objectively. Especially the continuous

use of closed-ended questions within the questionnaire left no room for possible interpretation.

Secondly, data harmonisation was counter-checked by a second person in order to reduce coding faults. This procedure was chosen because the previously described questionnaire design required entering numeric values at some points. The data harmonisation process is described in more detail in chapter 6.1.

INTERPRETATION OBJECTIVITY

Interpretation objectivity refers to the comparability of the conclusions inferred from the results of a survey by different interpreters (Rammstedt, 2004). Interpretation objectivity is therefore about definition and documentation of clear interpretation rules of the received data (Moosbrugger & Kelava, 2012).

Due to the quantitative character of this study, mainly quantitative approaches, such as correlation and regression analysis, were applied. Interpretations of proved and disproved relationships found can be followed by their numeric values. However, the author is clear on the responsibility that also numeric relationships must undergo examination of reasonability.

In sum, the objectivity of this research can be evaluated as comparatively high. The main reasons for this lie in the highly standardised questionnaire and the standardised data collection process. In addition, due to the previously mentioned further measurements, like counter-checking of data harmonisation, the objectivity criterion should be fulfilled.

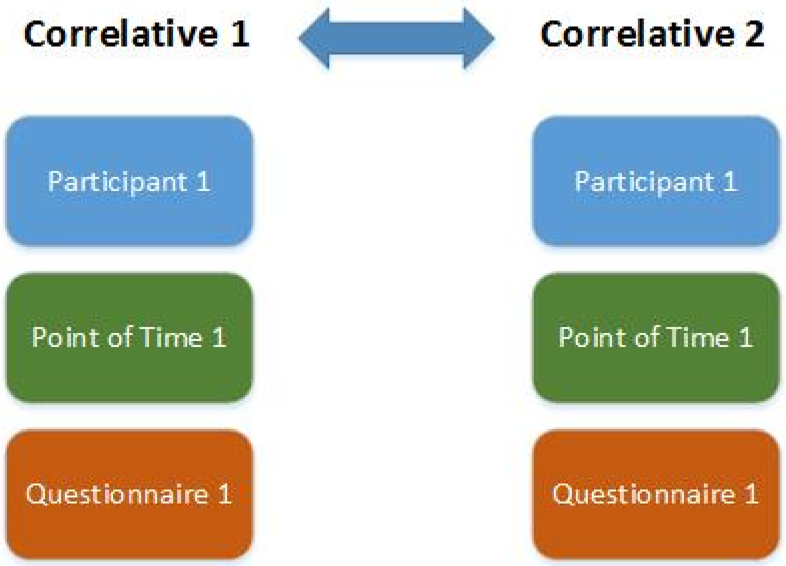
5.3.2 RELIABILITY

The second quality criterion is reliability. It refers to the extent to which data collection techniques or analysis procedures will yield consistency and stability, enabling findings to be replicated (Burns & Burns, 2008; Saunders et al., 2007).

In general terms, reliability can be understood as the degree of reproducibility of measurement results (Diekmann, 2004).

A measurement can be understood as reliable if the results remain constant when repeating the measurement. In order to verify reliability within this study, the survey would need to be replicated. Therefore, the same participants would have to answer the same questionnaire again at the same point of time. The results would be analysed for reliability of measurement by applying correlation techniques (Rammstedt, 2004). Figure 5.5 shows these circumstances in a general manner.

Figure 5.5: Theoretical concept of reliability assessment



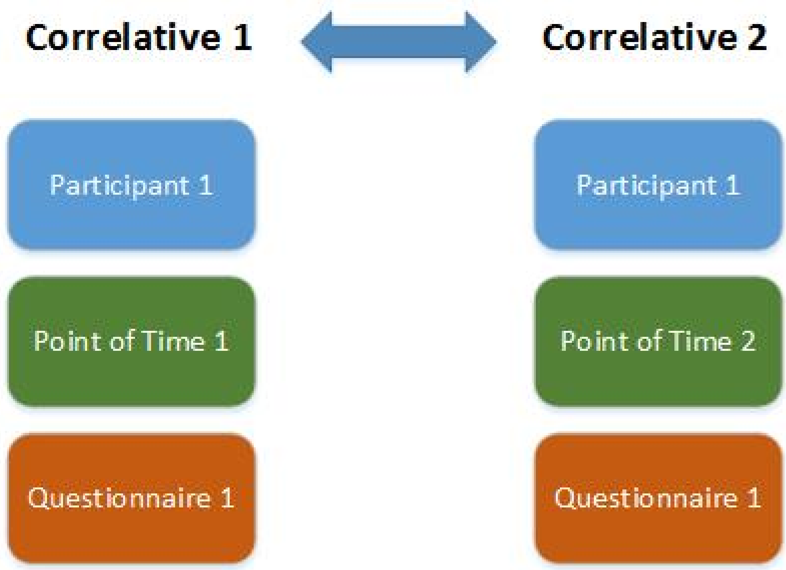
Source: Adapted from 5

However, such reliability assessments are often not feasible. Two completions of the same questionnaire at the same point of time are often not possible. One pragmatically application could be that just one participant repeats the survey just after completing the original one. But this could lead to memory effects which render this option questionable (Rammstedt, 2004)

Therefore, reliability must be assessed as a theoretical construct, which can only be evaluated by approximation (Rammstedt, 2004).
There are different methods to accomplish that. The most common reliability methods are "parallel reliability method", "split-half reliability method" and "internal consistency method", and "test-retest reliability method" (Burns & Burns, 2008; Diekmann, 2004; Rammstedt, 2004).

The author of this thesis assessed the test-retest method as being the most suitable technique to determine the reliability of the results.
Within the test-retest method, the questionnaire (or any other measuring instrument) will be re-conducted after a certain time interval (Diekmann, 2004). The correlation between the measurements from both points of time determines the test-retest reliability. Figure 5.6 shows a schematic view of the concept of test-retest reliability.

Figure 5.6: Concept of the test-retest reliability method



Source: Adapted from Rammstedt (2004), p. 6

In order to ensure that memory effects will not influence the second occasion, an adequate time period needs to be selected (Burns & Burns, 2008). According to Burns & Burns (2008), there is no standard duration between two occasions. They argue if the interval is too short, the participants could remember their answers, which would lead to an untrue increase of reliability. Aspects such as "boredom" or "decreased motivation" could also lead to a reduction of congruency between the two measurements. If the interval is too long, it could result in so-called "maturational factors", such as "learning experience" or "ageing of materials", which would have an effect on the values of the second occasion.

In this study, a complete retesting was not possible. The anonymous character of the survey made it difficult to re-contact the respondents.
In consequence, this study decided to apply a pragmatic approach and conduct the reliable test with the five companies interviewed in the pre-test.

The advantage of this procedure is that reliability could be determined before starting the survey. The disadvantage is, of course, that reliability is just tested on a small group and not for the whole sample size.

In any event, the five pre-test participants completed the questionnaire again three weeks after their first occasion. After data preparation, the correlation was calculated separately for each requested item. This enabled an overview of the reliability of every question/item/variable. According to Diekmann (2004), the reliability value should be at least 0.8.
In this study, each correlation achieved this result, with the average reliability amounting to 0.86.
This means that the criteria are fulfilled and reliability is given in this study.

However, objectivity and reliability of a measurement tool do not necessarily lead to valid measurements. Both are understood as minimum requirements of a measurement instrument (Diekmann, 2004).
Therefore, validity will be discussed in the following section.

5.3.3 VALIDITY

Burns & Burns (2008) highlight the importance of distinguishing between reliability and validity. They argue "whilst reliability relates to the accuracy and stability of a measure, validity relates to the appropriateness of the measure to assess the construct it purports to measure" (Burns & Burns, 2008, p. 425).
In corresponding literature, validity is often subdivided into internal and external validity.

EXTERNAL VALIDITY
External validity is defined as "the extent to which the results of a sample are transferable to a population" (Burns & Burns, 2008, p. 426). Sometimes external validity is therefore referred to as 'generalisability' (Saunders, et al., 2007). The level of this generalisation depends on the representation of the sample (population validity) and the natural conditions (ecological validity) (Burns & Burns, 2008; Gliner et al., 2009).

Population validity is defined as the extent to which the actual sample of participants represents the target population (Burns & Burns, 2008; Gliner et al., 2009).
Section 6.2.5 mentioned the application of a self-selection sample instead of a random sample within this study. In consequence, population validity is bound to have some weaknesses.

Ecological validity is mainly an issue of experimental research designs (Burns & Burns, 2008, p. 427). Since a non-experimental research design was selected for this study, ecological validity is assumed to not be relevant for this thesis.

INTERNAL VALIDITY
Internal validity instead regards the extent to which "any differences or relationships can be ascribed to the independent variable and not to any other factor" (Burns & Burns, 2008, p. 427). This means that internal validity is concerned with the question if a measurement instrument is actually measuring that for which it was intended. Internal validity generally comprises the following three aspects: content validity, criterion-related validity (predictive and concurrent validity) and construct validity (Burns & Burns, 2008; Creswell, 2009; Diekmann, 2004). These aspects will now be discussed with regard to the research design applied here.

Content validity is "the extent to which the content of a measurement reflects the intended content to be investigated" (Burns & Burns, 2008, p. 427). This can be tested e.g. through a previous debate with experts (Saunders et al., 2007).
In this research, the questionnaire and its appropriate covering of the research intent was discussed with five companies prior to the inquiry in the pre-test.

Predictive and concurrent validity are very similar and examine if the current measure is able to predict future performance (Burns & Burns, 2008). According to Burns & Burns (2008, p. 429), they can only by evaluated by "comparing a later performance (perhaps several years later) with the original performance source. It is usual to express predictive validity in terms of the correlation coefficient between the predicted status and the outcome criterion".

The survey was only taken once so far. To conduct it a second time would be an issue for further research and goes beyond the scope of this thesis. The predictive validity cannot be assessed in this research, unfortunately.

Construct validity means the level to which "the types of participant responses matches the intended construct" (Gliner et al., 2009, p. 167). This validity type should be considered when inquiring constructs like attitude scales (Saunders et al., 2007) and can be checked by questioning participants' intentions to give a certain answer (Gliner et al., 2009).

This proposal was followed by discussing the answers to the scale-related independent variables within the five-company pre-test.

This chapter showed that lots of measurements were realised to achieve good validity results in this thesis. The biggest limitation is probably population validity and consequently the generalisation of the further findings. The results have to be regarded very carefully. Moreover, this study is not able to assess its predictive validity.

Once a survey design has been chosen, it is important to think through the interaction with the participant. This is presented in the next chapter.

5.4 ETHICAL CONSIDERATIONS

Ethics play an important role when accessing, collecting, analysing, and reporting data from organisations and their individuals (Saunders et al., 2007). Burns & Burns (2008, p. 29) define it generally as "the application of moral principles and/or ethical standards that guide our behaviour in human relationships".

Several ethics codes and guidelines from different associations, such as NHS Research Governance Framework, British Sociological Association and Market Research Society, have been created to provide principles and procedures giving orientation on how to deal with this sensitive issue.

Easterby-Smith et al. (2008) report the work of Bell and Bryman (2007) who analysed the content of ethics codes from nine different associations and summarised them as 10 key principles of ethics presented in table 5.12.

Table 5.12: Key principles of research ethics

Number	Principle
1	Ensuring that no harm comes to participants
2	Respecting the dignity of research participants
3	Ensuring a fully informed consent of research participants
4	Protecting the privacy of research subjects
5	Ensuring the confidentiality of research data
6	Protecting the anonymity of individuals or organisations
7	Avoiding deception about the nature or aims of the research
8	Declaration of affiliations, funding sources, and conflicts of interest
9	Honesty and transparency in communicating about the research
10	Avoidance of any misleading, or false reporting of research findings

Source: (E. Bell & Bryman, 2007, cited in Easterby-Smith et al, 2008)

The author considers these principles a very good guideline for the interaction with his participants. Hence, these 10 key principles are followed in this research.

Topic 1 and 2 seem to be more important in medical, clinical or biometric research. Ensuring a safe and dignified participation in this context is more about respect, privacy and confidence.

The participation was voluntary and anonymous. Any dependence between the participants and the data set was impossible. At the end of the survey respondents had the opportunity to give personal details on a voluntary basis. This data was used in a strictly confidential manner and was only used for call-backs, if necessary.

Furthermore, the participants had been guaranteed that gathered data would only be used for the purpose of this research and not transferred to anyone else. This was stated in the invitation email.

In order to protect the participants from any inferences between the data and the person, no entire data sets will be published.

This leads also to the following of principle 4 and 6.

This thesis is not supported by any interest groups or stakeholders. Its findings therefore do not conflict with any other interests. The research aim is to contribute to a deeper understanding of ERP costs, which is a matter of public interest.

Reporting the findings transparently, honestly and neutrally is a basic interest of the research itself.
In addition, ethical considerations of this research have been approved by the University.

5.5 SUMMARY

This methodological review explains why the chosen paradigms, methods, designs and techniques were selected. The following list summarises the results:

- Paradigm:	Postpositivist paradigm
- Research design:	Quantitative approach
- Strategy of inquiry:	Associational approach
- Survey design:	Internet survey provides self-completion questionnaires
- Variables:	Either quantified by their nature or via Likert scale
	Variables are interval or ratio scales
- Sampling strategy:	Self-selecting sampling
- Data Collection:	EFS Survey
	Pilot Test, Pre-Survey Contact, 2 Reminders

6. DATA ANALYSIS & RESULTS

This chapter describes the data analysis and presents the findings of this study.
It begins by explaining the data preparation which is the basis for all further analyses.
The next step is about answering the research questions, presenting the findings and describing the conducted analysis.

6.1 DATA PREPARATION

After closing the survey, the gathered data needs to be prepared and harmonised. The data preparation process in this study consists of four steps:

1. Harmonising and coding data into appropriate formats & checking data for errors
2. Making the respondents' ERP costs comparable by discounting, respectively compounding their costs data
3. Transferring the data from Excel to SPSS
4. Checking the independent variables (x1 to x36) for multicollinearity among each other

These four steps are explained in more detail in the following:

At first, the collected data were exported from the EFS survey to Excel for further preparation. Since respondents should not be annoyed by prescribed data input formats during the survey, every single data needed to be checked and harmonised into one pre-defined standard format. To give some examples:
The maintenance costs of the lifecycle should be measured for a period of 10 years. Since not all participants operate their systems over that period of time and, of course, it is hard to give a precise answer to that question, the questionnaire asked for a period of one year. The answers then needed to be projected.
Other examples: Stated costs in "20k" needed to be changed into "20,000". Or the answers to the question "in which industry sector the company operates" needed to be categorised and summarised. Afterwards, the data was manually coded.

Furthermore, the data was checked for errors and corrected. In terms of contentual ambiguousness, respondents could be queried for clarification since most respondents voluntarily stated their email addresses. This was done in some cases. Afterwards the data was checked for eligibility and removed if necessary. The remaining responses were analysed for evidently implausible answers.
Every single response was thus checked in detail for unrealistic statements. In case of apparently unrealistic answers, the response was removed.

After this procedure had been completed, the actual sample size considered in this study included 72 valid responses. The next chapter views the respondents and the response rate in more detail.

The second step was to make the different responses comparable.
There were two main aspects to consider: Since the ERP projects of the respondents took place at different points in time, the first aspect was to regard the monetary inflation to make the costs comparable. The second aspect was to find an appropriate conversion rate for transferring the given costs in person-days into costs in EURO, because the questionnaire allowed both statements in order to facilitate its completion.
Considering the different points in time, the monetary inflation would indeed have had an impact on the cost values. Therefore, it was necessary to compound interest for the spent ERP costs to the value as of today. For that reason, all cost variables (y1-y21) were accumulated by an average inflation rate of 1.7% per year. Depending on the implementation year, variables were calculated correspondingly.

Transferring the efforts in person-days into the costs in EURO raises the question of a suitable conversion rate. Fortunately, several respondents stated the exact costs of their required internal personnel and external personnel person-days, which enabled the calculation of a realistic standard daily rate for internal as well as external personnel costs.

This study comes to the result that the internal and external personnel costs vary from one lifecycle phase to another.
Rounded, this study evaluates the internal personnel costs during the different lifecycle phases as follows:

EVALUATION: 600 €/PD
IMPLEMENTATION: 500 €/PD
MAINTENANCE: 600 €/PD

The costs of external personnel also vary between the phases. The rounded values are presented below:
EVALUATION: 1,000 €/PD
IMPLEMENTATION: 1,100 €/PD
MAINTANACE: 1,100 €/PD

These values represent the conversion rate from person-days into EURO within this study.

The third step of preparing data was shifting the data from Excel into SPSS and assigning the variables to their levels of numerical measurement, which is nominal, ordinal, interval, and ratio[16] data type in SPSS.
All y variables and most x variables are measured by quantifiable data, which means they can be positioned on a numerical scale and are either interval or ratio data types. The only exceptions are variable x5 and variable x12, which ask for the brand of ERP system and the type of implemented modules. These variables are measured in categories and must be defined as nominal data.

Unfortunately nominal data is actually not supposed to be used in statistical inferences. Since the relationship between the type of ERP system and the costs as well as the relationship between the types of modules and the costs should be examined, these variables needed to be prepared for correlation analysis by using the option to transfer them into so-called "dummy variables". Dummy variables can convert nominally scaled variables into binary variables and consequently handle them as metric variables (Backhaus, Erichson, Plinke, & Weiber, 2008).
The conversion process is as follows: Each possible specification of the variable gets its own variable. So variable x5 got 6 sub-variables by indexing its specifications as follows:

[16] SPSS does not distinguish between interval and ratio scales. SPSS summarises these two metric levels of measurement into the so-called "scale" level.

x5-1: SAP

x5-2: Microsoft

x5-3: proALPHA

x5-4: SAGE

x5-5: ABAS

x5-6: Other

Possible value parameters were 0 and 1.

Variable x12 (type of modules) was handled simultaneously and got 12 sub-variables, one for each specification. They were labelled x12-1 to x12-12.

The fourth and last step of the data preparation procedure is to prepare the data for the statistical inferences. For multiple correlation and regression analyses, it is important to consider the possibility of multicollinearity. In some studies, this step is taken as the first step during the actual regression, but this study aimed to have a valid data set at the end of the data preparation process. However, according to Burns & Burns (2008, p. 386), a very high correlation between independent variables implies that they "are measuring the same variance and will over-inflate R." Consequently, just one of them is needed. Very high correlations are values from 0.9 and above (Burns & Burns, 2008).

To check the variables x1 – x35 for their multicollinearity, a correlation matrix of all independent variables was generated.

Correlation coefficients over 0.9 are found only between the variable x3 (no. of total employees) and x10 (no. of total users). Since they measure the same variance, one of them can be disregarded in the multiple correlation and regression analysis.

The author decided to remove x3.

6.2 RQ 1: ANALYSIS & FINDINGS

Research Question 1 aims to find an answer regarding the quantity of incurred costs for ERP systems during their lifecycle phases and for their whole lifecycle.

The systematic literature review identified that the consideration of ERP costs has always been relatively fragmented, and almost every researcher focuses on different cost types.

The literature review identified seven different cost types. Within the conceptual framework, they were synthesised into one concept for the first time.

The table below briefly repeats them.

Table 6.1: Cost fields of the ERP lifecycle

	Internal Personnel Costs	External Personnel Costs	ERP Software Costs	Licence Costs	Hardware Costs
Evaluation	y1	y2	-	-	-
Implementation	y3	y4	y5	y6	y7
Maintenance	y8	y9	y10	y11	y12

The author of this thesis assumes that the costs for the whole ERP lifespan are composed of these five cost types during the three lifecycle phases evaluation, implementation and maintenance.

These cost types need to be assessed first in order to be able to give an answer regarding the costs of each phase and the whole lifecycle. Regarding the cost fields during each lifecycle phase, RQ1 is divided according to them and structured into several sub-questions which are repeated in the following:

Table 6.2: Research question I

RQ1: What are the costs of ERP systems during their lifecycle phases?	**RQ1-1**: What are the costs of internal personnel (y1) and external personnel (y2) during the evaluation lifecycle phase? **RQ1-2**: What are the costs of internal personnel (y3), external personnel (y4), ERP software (y5), licence (y6) and hardware (y7) during the implementation lifecycle phase? **RQ1-3**: What are the costs of internal personnel (y8), external personnel (y9), ERP software (y10), licence (y11) and hardware (y12) during the ERP lifecycle maintenance phase? **RQ 1-4:** What are the costs of ERP systems during their whole lifespan?

Each sub-question of RQ1 will be regarded in its own sub-chapter.
The results are generated by average and average absolute deviation (AAD) in absolute and relative terms.

6.2.1 RQ1-1: ANALYSIS & FINDINGS

First the data analysis regarded the absolute expenses for internal (y1) and external personnel costs (y2) during the evaluation phase.
The table below gives an overview of the average absolute costs, the average absolute deviation for internal personnel and external personnel costs during the evaluation phase as well as the minimum and maximum expenses for these cost fields within the data set.
Total evaluation regards the cumulative costs for the whole evaluation phase and describes the cumulative minimum and maximum costs of the data set.

Table 6.3: Absolute costs for y1 and y2 during evaluation phase

	Internal Personnel Costs (y1)	External Personnel Costs (y2)	Total Evaluation
Average	35,855 €	30,307 €	66,163 €
AAD	31,129 €	43,659 €	69,456 €
SD	49,117 €	96,111 €	138,985 €
Min.	0 €	0 €	0 €
Max.	305,100 €	711,900 €	1,017,000 €

The table shows that, on average, the companies in this sample paid 35,855 EUR for internal personnel and 30,307 EUR for external personnel. In total, they spent 66,163 EUR for the whole evaluation phase.

However, regarding the absolute average deviation (AAD), it shows that the values are dispersed widely around the mean. That means that the values spread widely and the expenses vary strongly within the sample.
Looking at the minimum and maximum expenses within the sample shows that the expenses for internal personnel vary between 0 to 305,855 EUR and for external personnel between 0 EUR and 711,900 EUR. The costs for the whole evaluation phase range from 0 to 1,017,000 EUR.

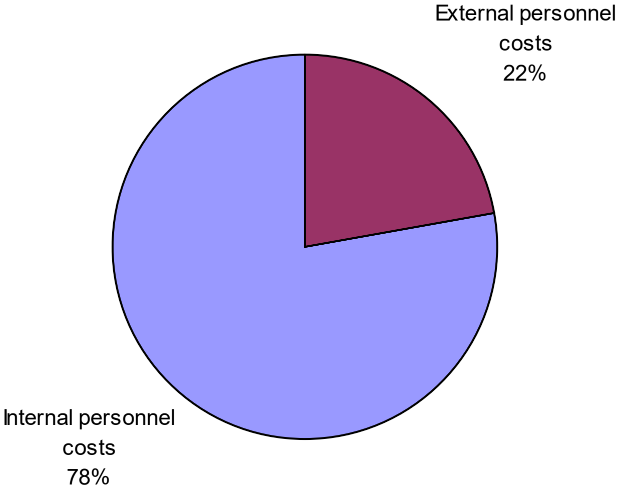
Regarding the cost distribution during the evaluation phase, analysis shows that companies incur a lot more costs for internal than for external personnel.
Table 6.4 presents the costs of internal and external personnel during the evaluation phase expressed as percentages. The percentages of minimum and maximum expenses present the lowest and highest costs incurred within the sample as a relative value.

Table 6.4: Percentage costs for y1 and y2 during the evaluation phase

	Internal Personnel Costs (y1) in %	**External Personnel Costs (y2) in %**
Average	77.7%	22.3%
SD	25.1%	25.1%
Min.	0 %	0 %
Max.	100%	83.3%

On average, the costs during the evaluation phase are divided into 78% for internal personnel and 22% for external personnel, as shown in the table. The costs for internal personnel are thus 3.5 times higher than for external personnel. The pie chart below highlights this distribution.

Figure 6.1: Average ratio of y1 and y2 during ERP evaluation phase



However, table 6.4 also implies that the cost distribution varies strongly within the data set. The value 100% shows that at least one company has just expenses for internal personnel and no costs for external staff.

83.3% in the column External Personnel Costs means that the percentage of maximum costs of one company was 83.3% of its total costs during the evaluation phase. This value differs extremely from the central tendency.

Being confronted with these varying values, the percentage average cost structure should be identified. It might be helpful to know how many of the whole lifecycle costs are required for the evaluation phase to make a more precise statement. This is determined by cost ranges which define the average variation of costs within these cost fields.

Therefore, the percentage cost distribution of the cost fields "internal personnel costs" (y1) and "external personnel costs" (y2) as well as their Average Deviation (AD) need to be calculated:

$$AD = \frac{1}{N} \sum \left| y_i\% - \bar{y}\% \right|$$

The results are presented below:

Table 6.5: Cost ranges of cost fields in evaluation phase

Cost Field during Evaluation Phase	Average Percentage of Total Lifecycle Costs	AD	Percentage Cost Range of Total Lifecycle Costs
Internal Personnel costs (y1)	3.7%	+/-3.0%	0.7 – 6.7%
External Personnel costs (y2)	2.1%	+/-2.6%	0.0 – 5.2%
Total Evaluation	5.8%	+/-8.0%	0.0 – 16.0%

The analysis identified that the cost range for the whole evaluation phase was between 0 and 16% (average 5.8%) of the whole lifecycle costs on average.
The portion for internal personnel costs (y1) varies on average between 0.7% and 6.7% (average 3.7%). External personnel costs require between 0 and 5.2% (average 2.1%) of the budget.

To conclude this section, the most important results of the evaluation phase are briefly summarised:

Internal personnel costs (y1)

The average absolute expense for internal personnel is 35,855 €. This cost field accounts for 77.7% of the total costs during the evaluation phase.

However, the expense varies strongly among the sample, namely between 0 € and 305,100 € in absolute terms. Regarding the percentage value, at least one company incurs all costs during its evaluation phase for internal personnel.

To make a more precise statement, the average percentage cost ranges of each cost field are determined in dependence on the whole ERP lifecycle costs. The result is that, on average, the costs amount to 3.7% and vary between 0.7 and 6.7% of the whole ERP lifecycle costs.

External personnel costs (y2)

On average, the costs for external personnel during the evaluation phase amount to 30,307 €. 22.3% of expenses during the evaluation are incurred for external staff, on average. Nevertheless, the costs strongly vary between 0 € and 711,900 € in absolute terms. The percentage value shows that at least one sample case incurs all costs during its evaluation phase for external staff.

The average cost distribution of external personnel costs during the whole ERP lifecycle is 2.1% and ranges between 0 and 5.2%.

Evaluation phase

On average, 66,163 € are spent for the evaluation lifecycle phase. The minimum expense in the data set is 0 € and the maximum cost is 1,017,000 €. With regard to the costs of the whole ERP lifecycle, the evaluation phase requires on average 5.8% and the costs range between 0 and 16%.

6.2.2 RQ1-2: ANALYSIS & FINDINGS

The analysis of RQ1-2 finds that ERP software costs create the highest cost factor during the implementation phase.

Table 6.6 shows that software costs (y5) require the highest expenditure: This cost field needs 136,953 EUR on average. With averagely 136,637 EUR, it is followed by costs for internal personnel (y3).

With 112,320 EUR, external personnel costs (y4) are on the third place. The fourth place goes to licence costs (y6) and the last one goes to hardware costs (y7). In total, the implementation phase costs 543,613 EUR on average.

Table 6.6 gives an overview of the average costs for each cost field during the implementation phase, their average absolute deviation as well as the minimum and maximum expense.

Table 6.6: Absolute costs for y3 – y7 during the implementation phase

	Internal Personnel Costs (y3)	**External Personnel Costs (y4)**	**ERP Software Costs (y5)**	**Licence Costs (y6)**	**Hardware Costs (y7)**	**Total Costs for Implementation Phase**
Average	136,637 €	112,320 €	136,953 €	99,454 €	58,250 €	543,613 €
AD	132,041 €	112,031 €	122,402 €	90,993 €	57,093 €	419,398 €
SD	179,025 €	140,469 €	157,510 €	121,839 €	80,046 €	510,860 €
Min.	0 €	0 €	0 €	0 €	0 €	21,759 €
Max.	828,529 €	524,490 €	598,367 €	591,806 €	473,445 €	2,071,321 €

Besides the average values of expense grouped by cost fields, Table 7.7 presents a similar picture of strongly varying costs within the sample.

The cost field Internal Personnel Costs shows the biggest variations. It has the highest average absolute deviation, which is 132,041 EUR. When considering the lowest and the highest expense for this cost field within the sample, the costs vary between 0 and 828,529 EUR.

The external personnel costs range from 0 to 524,490 EUR, ERP software costs from 0 to 591,806 EUR, licence costs from 0 to 591,806 EUR and the hardware costs from 0 to 473,445 EUR.

In the next step, the percentage cost distribution during the implementation phase should be assessed. Table 6.7 presents the average percentage distribution as well as the percentage minimum and maximum expense within the sample. The percentage minimum and maximum expense present the lowest and highest costs incurred within the sample in a relative value.

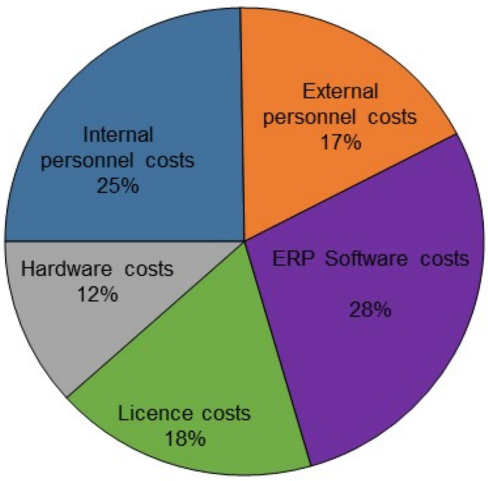
Table 6.7: Percentage costs for y3 – y7 during the implementation phase

	Internal Personnel Costs (y3) in %	External Personnel Costs (y4) in %	ERP Software Costs (y5) in %	ERP Licence Costs (y6) in %	ERP Hardware Costs (y7) in %
Average	24.7%	17.6%	28.1%	17.8%	11.7%
SD	17.7%	14.4%	19.9%	14.4%	10.1
Min.	0%	0%	0%	0%	0%
Max.	100.0%	79.5%	78.4%	63.3%	48.9%

The distribution of cost fields during ERP implementation shows that ERP software costs (y5) amount to 28% of the total implementation costs. This value is close to the internal personnel costs (y3), representing 25% of the costs.
External personnel costs (y4) and licence costs (y6) each amount to 18%. Hardware costs (y7) with 12% of the total implementation costs make up the smallest portion.

Figure 6.2 graphically illustrates this cost structure.

Figure 6.2: Average ratio of ERP implementation costs



At this stage again, the expense of the different cost fields differs highly from company to company. This can also be seen when considering the minimum and maximum value in percentages for each cost type presented in table 6.7.
The maximum value of 100% for the internal personnel costs means that at least one company had expenses just for that cost field and did not spend anything on external staff, software, licence or hardware.

Although the other maximum values do not reach a full 100%, their values are still relatively high.

Nevertheless, when comparing the cost fields Internal and External Personnel Costs in absolute terms to the ones of the evaluation phase, it can be seen that the costs for internal personnel costs are 3.8 times higher and the ones for external personnel are 3.7 times higher than during the evaluation phase.

Finally, it should be considered how many of the overall lifecycle costs arise during the implementation phase.
The percentage cost distribution of the cost fields “internal personnel costs” (y3), “external personnel costs” (y4), “ERP software costs” (y5), “licence costs” (y6) and “hardware costs” (y7) thus need to be calculated in dependence on the whole arising costs:

$$AD = \frac{1}{N}\sum \left| y_i\% - \bar{y}\% \right|$$

The result is presented in the table below.

Table 6.8: Cost Ranges of cost fields during the implementation phase

	Average	Average Deviation	Cost Range
Internal Personnel Costs (y3)	9.8%	+/- 6.6%	3.2 - 15.8%
External Personnel Costs (y4)	7.0%	+/- 5.0%	2.0 – 12.0%
ERP Software Costs (y5)	12.5%	+/- 10.1%	2.4 – 22.6%
Licence Costs (y6)	7.0%	+/- 5.1%	1.9 – 12.1%
Hardware Costs (y7)	4.8%	+/- 3.7%	1.1 – 8.5%
Implementation Phase Total	**41.1%**	**+/- 12.0%**	**29.1 – 53.1%**

The cost range for the whole implementation phase varies between 29.1 and 53.1% of the whole ERP lifecycle costs on average.
Internal personnel costs require between 3.2 and 15.8% (average 9.8%) and external personnel costs range from 2.0 to 12.0% (average 7%) on average. The biggest range can be discovered in the software costs which range from 2.0 to 22.6% (average 10.1%). Licence

costs vary between 1.9 and 12.1% (average 4.8%) on average. The smallest range is discovered for hardware costs which differ between 1.1 and 8.5% (average 4.8%).

To conclude this section, the most important results of the implementation phase are briefly summarised:

Internal Personnel Costs (y3)

Internal personnel costs require the second largest portion of costs arising during the implementation phase. In absolute terms, 136,637 EUR are spent in this cost field. Expressed in relative values, this is 24.7%.

The expense within the data set varies strongly, namely between 0 € and 828,529 €. The sample case with the highest personnel costs spent 100% in this cost field, which means that this case did not have any costs for software, licence, etc.

To make a more precise statement, the average percentage cost ranges of each cost field are determined in dependence on the whole ERP lifecycle costs. They are 9.8% on average and range from 3.2% to 15.8%.

External Personnel Costs (y4)

On average, the costs for external personnel are 112,320 € during the evaluation phase. This corresponds to the third highest expense during the implementation phase.

Regarding the ratio of costs arising during the implementation, 17.6% account for this cost field. The absolute expense varies between 0 € and 524,490 €.

The percentage value shows that at least one sample case spent 79.8% of all costs during implementation for external staff.

The average cost distribution for external personnel costs during the whole ERP lifecycle is 7.0% and ranges from 2.0 to 12.0%.

ERP Software Costs (y5)

This cost field requires the biggest budget during the implementation phase. 136,953 EUR are averagely spent for ERP software. With regard to the ratio of all costs incurred during the implementation phase, this means 28.1%.

The expenses differ from a minimum of 0 EUR to a maximum of 598,367 EUR in absolute terms. Of the sample, the company with the highest relative expenses in this cost field shows a value 78.4%.

Depending on the whole ERP lifecycle costs, the range of ERP software costs during the implementation phase is between 2.4% and 22.6%, and 12.5% on average.

Licence Costs (y6)

The licence costs are on average 99,454 EUR. They have an average ratio of 17.8%.
In absolute terms, the costs for this cost field vary between 0 and 598,367 EUR. The relative maximum expense is 63.3%
With regard to the whole lifecycle, the average percentage cost distribution is 7.0%. The cost range for licence costs during the whole ERP lifecycle is 1.9% to 12.1% on average.

Hardware Costs (y7)

With 58,250 EUR on average, this cost field is the one with the lowest expense. Expressed in relative values, 11.7 % of the arising costs are averagely spent for hardware during this lifecycle phase.
Nevertheless, its difference of minimum and maximum costs varies strongly between 0 and 473,445 EUR in the sample. The sample case with the highest hardware costs spent 48.9% for this cost field.
Regarding the costs in dependence on the whole ERP lifecycle, the average hardware costs account for 4.8% and vary between 1.1% and 8.5%.

Implementation Phase

543,613 € are averagely spent during the implementation lifecycle phase. The minimum expense in the data set is 21,759 €, whereas the maximum cost is 2,071,321 €. Regarding the cost range for the whole ERP lifecycle, the implementation phase requires 41.1% and varies between 29.1% and 53.1% of the whole ERP lifecycle costs on average.

6.2.3 RQ1-3: ANALYSIS & FINDINGS

ERP maintenance costs are incurred costs which repeat annually. This study assumes that ERP systems have a durability of 10 years after their implementation, as mentioned earlier. Since not all inquired companies keep their systems for 10 years and it is difficult to estimate one’s maintenance costs over a period of 10 years, the questionnaire asked for the annual maintenance costs. This means that the stated costs needed to be multiplied by 10.

Table 6.9 regards the average costs for each cost field in absolute terms during a 10-year maintenance phase. Besides that, it shows the average absolute division and the minimum and maximum expense within the sample.

Table 6.9: Costs for y8 – y12 during maintenance phase in absolute terms

	Internal Personnel Costs y8	External Personnel Costs y9	ERP Software Costs y10	ERP Licence Costs y11	ERP Hardware Costs y12	Total Maintenance Costs
Average	390,280 €	190,620 €	83,140 €	225,550 €	95,830 €	985,430 €
AD	500,400 €	216,320 €	103,710 €	193,610 €	10,361 €	1,024,401 €
SD	760,554 €	290,079 €	170,430 €	245,700 €	172,010 €	1,306,600 €
Min.	0 €	0 €	0 €	0 €	0 €	0 €
Max.	4,000,000 €	1,250,000 €	1,000,000 €	960,000 €	1,200,000 €	8,410,000 €

The analysis finds that, in absolute terms, internal personnel costs (y8) require the biggest budget during a 10-year maintenance phase.
The costs of 390,280 EUR for this cost field are followed by licence costs (y11) with 255,550 and expenses in the amount of 190,620 EUR for external personnel costs (y9). With 95,830 EUR, the second last position goes to hardware costs (y12), followed by software costs (y9) in the amount of 83,140 EUR.

However, also at this point, the data dispersion is relatively high, which means that the values spread widely around the average expense. The minimum and maximum expenses vary significantly within the sample. The highest variation is in the costs for internal personnel (y8), which differ between 0 and 4,000,000 EUR within the sample. The expense for licence costs (y11) shows the least difference of between 0 and 960,000 EUR.

Looking at the costs in relative terms creates a different picture. Table 6.10 gives an overview of the average percentage cost distribution during 10 years of maintenance.
The percentage of minimum and maximum expense represents the lowest and highest costs incurred within the sample in a relative value.

Table 6.10: Costs for y8 – y12 during maintenance phase in relative terms

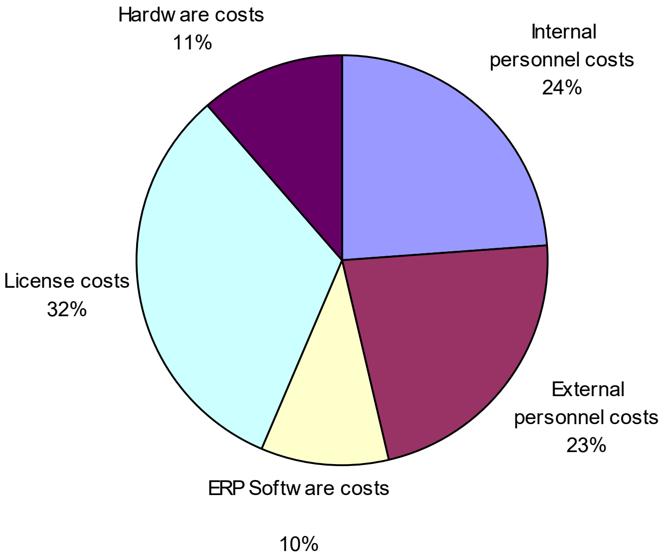
	Internal Personnel Costs in %	External Personnel Costs in %	ERP Software Costs in %	ERP Licence Costs in %	ERP Hardware Costs in %
Average	23.8%	22.5%	10.1%	32.2%	11.5%
SD	25.9%	26.4%	18.9%	30.0%	13.8%
Min.	0%	0%	0%	0%	0%
Max.	88.2%	100.0%	94.2%	100.0%	73.5%

In contrast to the absolute cost distribution, the relative one finds that the highest relative expense is for licence costs (y12). This cost field requires averagely 32% of the total costs during the maintenance phase.

Licence costs are followed by internal personnel costs (y8) with averagely 23.8%. Requiring 22.5% on average, external personnel costs have the third place within the cost distribution during the maintenance phase. Hardware costs and software costs are on the last positions.

The pie chart below shows the cost distribution of each cost field during the maintenance phase:

Figure 6.3: Average ratio of y8 – y 12 during the maintenance phase



Finally, the question how many of the overall lifecycle costs arise during the maintenance phase should be answered.

To this end, the percentage cost distribution of the cost fields "internal personnel costs" (y8), "external personnel costs" (y9), "ERP software costs" (y10), "licence costs" (y11) and “hardware costs” (y12) need to be calculated in dependence on the whole costs incurred.

The result is presented below:

Table 6.11: Cost ranges of cost fields during maintenance phase

	Average	Average Deviation	Cost Range
Internal Personnel Costs (y8)	14.8%	+/- 15.1%	0.0 – 30.5%
External Personnel Costs (y9)	10.9%	+/- 9.5%	1.4 – 20.4%
ERP Software Costs (y10)	5.3%	+/- 5.9%	0.0 – 11.8%
Licence Costs (y11)	16.3%	+/- 11.5%	4.8 – 27.8%
Hardware Costs (y12)	5.9%	+/- 4.8%	1.1 – 10.7%
Total Maintenance Phase	**53.2%**	**+/- 9%**	**44.2- 62.2%**

The whole maintenance costs range from 44.2% to 62.2% (average 53.2%) of the whole lifecycle costs.

The range of internal personnel costs is between 0% and 30.5% (average 14.8%). External personnel costs range from 1.4% to 20.4% (average 10.9%) on average. Software costs require between 0% and 11.8% (average 5.3%) and licence costs between 4.8% and 27.8% (average 16.3%). The smallest range discovered is for hardware costs, which vary between 1.1% and 10.7% (average 5.9%).

To conclude this section, the most important results of the maintenance phase are briefly summarised:

Internal Personnel Costs (y8)

Regarded in absolute terms, this cost field requires the highest expense with averagely 390,280 EUR. Considering how big the portion of this cost field is in relation to all costs incurred during 10 years of maintenance, the value averages 23.8%. This is the second largest portion.

Nevertheless, the average value should not suppress the fact that both the highest and lowest absolute and relative values within the sample vary strongly from one company to another. The absolute values vary between 0 and 4,000,000 EUR.

The relative values vary between 0% and 88.2%. That means that at least one company within the sample spent 88.2% just for its internal personnel.
Analysing y8 in dependence on the costs of the whole lifecycle, the average cost is 14.8%. Its range is between 0% and 30.5%.

External Personnel Costs (y9)

External personnel costs require the third highest expense during the maintenance phase. In absolute terms, the costs average 190,620 EUR. When regarding them in dependence on all costs incurred during this phase, this cost field requires 22.5%.
The highest and lowest absolute and relative values within the sample differ: The absolute values between 0 and 1,250,000 EUR and the relative values between 0% and 100%. That means that at least one company only had expenses for external staff during its maintenance phase.
The average cost range of external personnel costs in the maintenance phase is 10.9%. It ranges from 1.4% to 20.4% when considering the expenses incurred for the whole lifecycle.

ERP Software Costs (y10)

83,140 EUR are averagely spent for ERP software costs during the 10-year maintenance phase. These are the lowest costs and correspond to a relative value of 10.1%. That means that 10.1% of all costs arising during this phase account for this cost field on average.
The expenses differ from a minimum of 0 EUR to a maximum of 1,000,000 EUR in absolute terms. Of the sample, the company with the highest relative expenses in this cost field shows a value of 94.2%.
In dependence on the whole ERP lifecycle costs, ERP software costs account for 5.3% and vary between 0% and 11.8% during the implementation phase.

Licence Costs (y11)

When considering the expense for each cost field in dependence on all costs arising during the maintenance phase, licence costs are the ones with the highest expenditure with 32.2% on average. In absolute terms, this cost field requires 225,550 on average, which is the second highest value, and ranges from 0 to 960,000 EUR within the sample. The proportionally highest expenses in the sample equal 100%. Regarding the licence costs during the maintenance phase in dependence on the costs of the whole ERP lifecycle results in an average cost of 16.3% and a range of 4.8% to 27.8%.

Hardware Costs (y12)

With expenses of averagely 95,830 €, the second lowest costs during the maintenance phase are in this cost field. Expressed in relative values, 11.5% of the costs incurred are averagely for hardware. However, the expense within the data set varies strongly, namely between 0 € and 1,200,00 €. The sample case with the highest personnel costs spent 73.5%.

Analysing y12 in dependence on the costs of the whole lifecycle, the average range of expense is 5.9% and varies between 1.1% and 10.7%

Maintenance Phase

985,430 € are averagely spent for the maintenance lifecycle phase. The minimum expense in the data set is 0 €, whereas the maximum cost is 8,410,000 €. Even if the value of 0 EUR seems to be very unlikely, the possibility cannot be excluded. The next lowest value in the data set is 13,200 EUR per year.

Nevertheless, regarding the costs for the whole ERP lifecycle, the evaluation phase requires averagely 53.2% and ranges from 44.2% to 62.2%

6.2.4 ERP COSTS DURING THE TOTAL LIFECYCLE

After analysing the cost fields during each lifecycle phase, the costs for the whole lifespan are regarded in the following. To this end, each cost field is cumulated with the cost fields of the other lifecycle phases.

The results are presented in table 6.12 below. Besides showing the average expenses in absolute terms, it presents the standard deviation and the minimum and maximum costs within the sample for every cost field.

Table 6.12: Costs for the cumulated cost fields during the whole ERP lifecycle

	Total Internal Personnel	Total External Personnel	Total ERP Software Costs	Total Licence Costs	Total Hardware Costs	Total Lifecyc-le
Average	562,777 €	333,250 €	220,092 €	325,004 €	154,083 €	1,595,206 €
AAD	609,685 €	325,309 €	188,726 €	255,888 €	146,663 €	1,526,271 €
SD	917,819 €	425,995 €	252,897 €	320,773 €	227,983 €	1,772,244 €
Min.	0 €	0 €	0 €	0 €	0 €	0 €
Max.	4,840,365 €	1,809,350 €	1,342,775 €	1,265,100 €	1,430,498 €	10,688,088

The analysis finds that, in absolute terms, internal personnel costs account for the highest expense during the whole lifespan.
The costs of 562,777 EUR in total for this cost field are followed by licence costs with 325,004 EUR and expenses in the amount of 333,250 EUR for external personnel costs. The second to last position goes to software costs with a cumulated amount of 220,092 EUR. The lowest total costs are hardware costs with 154,083 EUR in total for the whole ERP lifespan.

Unsurprisingly, here too, the average absolute deviation is high, and the expenses vary strongly within the sample. The biggest deviation was discovered within the cost field internal personnel costs for the whole lifespan. The absolute costs vary between 0 EUR and 4,840,365 EUR. At this stage, again, the value of 0 is extremely unlikely but cannot be excluded since the possibility exists.

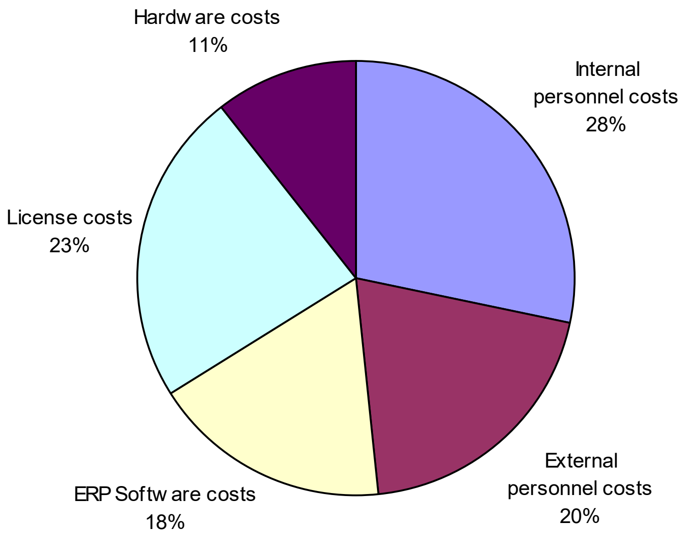
The costs of each cumulated cost field, depending on all costs arising during the whole lifespan, show the following results:

Table 6.13: Average percentage composition of cost fields during the lifecycle

	Internal Personnel Costs in %	External Personnel Costs in %	ERP Software Costs in %	ERP Licence Costs in %	ERP Hardware Costs in %
Average	28.3%	20.0%	17.8%	23.3%	10.6%
SD	18.5%	13.6%	14.5%	17.2%	10.1%
Min.	0%	0%	0%	0%	0%
Max.	72.3%	60.0%	74.5%	74.6%	52.4%

Table 6.13 presents that the major cost field, having a percentage of 28.3%, requires on average nearly one third of all costs incurred during a lifespan.
The second highest portion is licence costs with 23.3%, followed by external personnel costs amounting to 20%. With 17.8%, the fourth place goes to ERP software.
The lowest costs during the whole ERP lifespan arise for hardware with 10.6%.
The pie chart below illustrates this cost distribution:

Figure 6.4: Average ratio of cost fields during the whole lifecycle



Regarding the relative minimum and maximum expenses in table 6.13, the values high-light the different expenses for each cost field within the sample. This fact is considered in the next chapter, Results of RQ1.

6.2.5 SUMMARY & CONCLUSION RQ1

RQ1 managed to identify a percentage cost structure of cost fields for the costs arising during each ERP lifecycle phase and for the whole ERP lifespan. The major results are summarised in this chapter.

First, the analysis of RQ1 discovered the average expenses for each cost field during each lifecycle phase in absolute terms. The table below gives an overview of the average costs for each cost field within the sample:

Table 6.14: Average expenses of each cost field in EUR

	Evaluation	Implementation	Maintenance (10 years)	Whole Lifecycle
Internal Personnel	35,855	136,637	390,280	562,777
External Personnel	30,307	112,320	190,620	333,250
ERP Software	-	136,953	83,140	220,092
Licence	-	99,454	225,550	325,004
Hardware	-	58,250	95,830	154,083
Total	**66,163**	**543,613**	**985,430**	**1,595,206**

The extent to which the data values deviate from the average value in the previous subchapters was indicative of very wide value dispersions. The expenses vary strongly from one company to another.

The identification of the average costs gives a first impression, but has its limitations since these values cannot be generalised at all.

However, these absolute values enabled the generation of a percentage cost distribution of ERP costs during the whole lifespan.

This cost distribution, which can also be called cost structure, allows a more precise understanding of which costs arise at what time stage during the ERP lifespan.

Therefore it determined the average percentage quantity of each cost field in dependence on all costs arising during the whole ERP lifespan and its average percentage deviation. The so created cost ranges are able to show the average expenses of each cost field in a very detailed manner. The results are listed in the following table 6.15.

Table 6.15: Average percentage and its average percentage deviation

	Evaluation	Evaluation AD	Implementation	Impl. AD	Maintenance	Maintenance AD
Internal Personnel	3.7%	3.0%	9.8%	6.6%	14.8%	15.1%
External Personnel	2.1%	2.6%	7.0%	5.0%	10.9%	9.5%
ERP Software	-		12.5%	10.1%	5.3%	5.9%
Licence	-		7.0%	5.1%	16.3%	11.5%
Hardware	-		4.8%	3.7%	5.9%	4.8%
Total	**5.8%**	**8.0%**	**41.1%**	**12.0%**	**53.2%**	**9.0%**

This table allows the mapping of a more detailed picture of the cost structures in ERP lifecycles. It shows not only the average percentage distribution of costs, but also the cost ranges where the cost fields usually occur.
This provides added value since knowing the average range of expenses in dependence on whole costs is certainly helpful for project managers, implementing companies and ERP vendors in several ways.

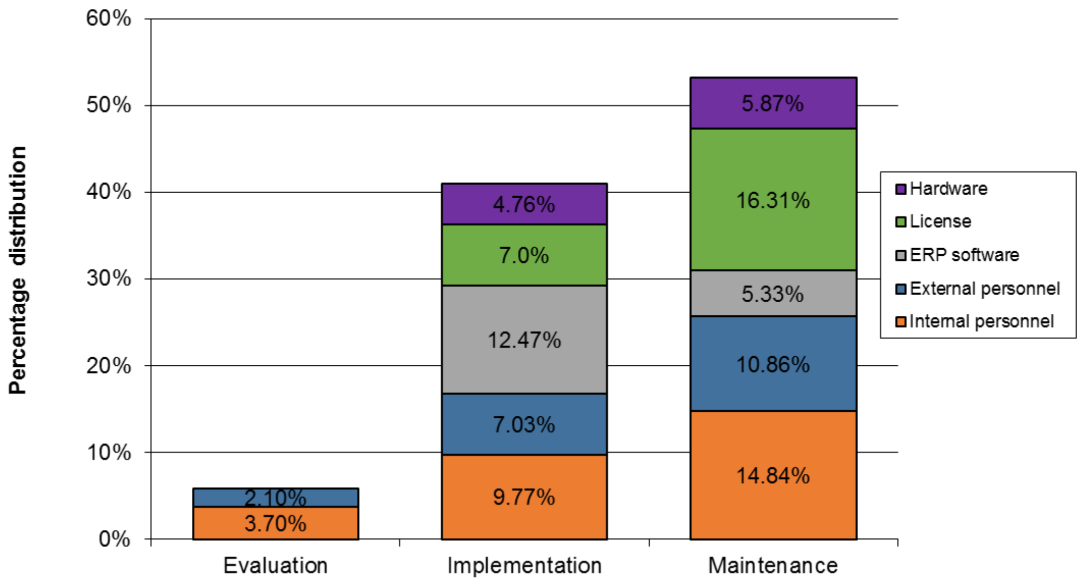
RQ1 identified that the maintenance phase (regarded for 10 years) averagely is the one with the most costs incurred. It requires 53.2% on average and ranges between 44.2% and 62.2% of the whole ERP lifecycle costs.
During this phase, the licence cost is the most dominant factor with 16.3% +/-11.5%. This value is followed by internal personnel (14.8% +/- 15.1%) and external personnel (10.9% +/- 9.5%). The lowest expenses arise for hardware (5.9% +/- 4.8%) and software (5.3% +/- 5.9%) at this stage.

By contrast, software costs are the cost field with the highest costs during the implementation phase, requiring 12.5% +/- 10.1% on average.
This expense is followed by internal personnel (9.8% +/- 6.6%). Both external personnel costs and licence costs are averagely 7.0%. The average deviation for external personnel is +/- 5.0% and for licence +/- 5.1%. The lowest costs arise for hardware (4.8% +/- 3.7%) during the evaluation phase.
In total, the implementation phase incurs on average 40.9% +/- 12% of the whole ERP lifecycle costs.

The lowest budget is needed for the evaluation phase. It requires 5.8% +/- 8% of the total costs. With 3.7% +/- 3.0%, internal personnel costs require on average a little more than the external personnel costs (2.1% +/-2.6%) of the total budget.

Notwithstanding the average deviation, figure 6.5 highlights the average percentage cost distribution of the whole ERP lifespan.

Figure 6.5: Average percentage cost distribution of the whole ERP lifecycle



These findings enable a very detailed insight into the average percentage cost distribution of the whole ERP lifecycle. These identifications create transparency about the costs arising over a time period of more than 10 years (evaluation + implementation + 10 years of maintenance) and provide further knowledge about ERP costs.

After having determined the percentage cost distribution, the next chapter aims to determine which cost drivers are de facto responsible for the quantity of costs arising within each cost field. It will concentrate on establishing relationships between cost drivers and the costs of each cost field.

6.3 RQ2: ANALYSIS & FINDINGS

After discovering a percentage cost structure by answering RQ1, RQ2 aims to assess which of the cost drivers are responsible for the extent of ERP expenditures.

As a first step, all ever stated cost drivers are cumulated into one concept. They are identified by the conducted systematic literature review and derived from former models, case studies or theoretical models in the existing ERP cost literature.
Most of the cost drivers have never been validated before. This chapter pools all of them.

Furthermore, none of the found literature regards cost drivers in a differentiated context. Many studies do not state which costs they relate to and at which time they arise. This study, by contrast, creates a clear connection between cost drivers, cost fields and the lifecycle phases they arise in. RQ2 thus concentrates on investigating relationships between cost drivers and the costs of each cost field and aims to measure their strength.

To this end, the cost drivers are investigated within each cost field (y1 – y12). This increases the transparency of the cost composition for the whole ERP, since the analysis might allow ascribing specific cost drivers to specific cost fields.

Chapter 4 presented an overview of all regarded cost drivers in this study. They are shown again in table 6.16.

Table 6.16: Overview cost drivers

Organisational Cost Drivers	Technical Cost Drivers	Situational Cost Drivers
x1: No. of locations	x5: ERP system	x16: No. of consultants
x2: No. of organisational units or depts.	x6: No. of interfaces	x17: No. of project members
x3: No. of total employees	x7: No. of modifications	x18: Ratio external / internal
x4: Revenue	x8: No. of reports	x19: Fit ERP system / organisation
	x9: No. of EDIs	x20: Team quality
	x10: No. of total users	x21: Team maturity
	x11: No. of user groups	x22: Team composition
	x12: Type of modules	x23: Availability of management
	x13: Complexity of data	x24: Availability of business users
	x14: Complexity of interfaces	x25: Consulting experience
	x15: Complexity of reports	x26: Consulting quality
		x27: Critical attitude of users
		x28: User quality
		x29: Employee involvement
		x30: Management involvement
		x31: Maturity of processes
		x32: Complexity of business processes
		x33: Stability of organisation
		x34: Willingness to change
		x35: Motivation of implementation team

As in the previous chapter, the cost fields are considered for each lifecycle phase. Consequently, RQ2 is divided into the corresponding sub-questions which are repeated in the following table:

Table 6.17: Research question II

RQ2: Which cost drivers influence ERP costs?	**RQ2-1:** Which cost drivers influence internal personnel costs (y1) and external personnel costs (y2) in the evaluation lifecycle phase?
	RQ2-2: Which cost drivers influence internal personnel costs (y3), external personnel costs (y4), ERP software costs (y5), licence costs (y6) and hardware costs (y7) in the implementation lifecycle phase?
	RQ2-3: Which cost drivers influence internal personnel costs (y8), external personnel costs (y9), ERP software costs (y10), licence costs (y11) and hardware costs (y12) in the maintenance lifecycle phase?

Aiming to measure the relationship between cost drivers and cost fields, these research questions should be answered by conducting a correlation analysis for each cost field.

For ordinal and interval data, the most widely used coefficient is the so-called 'Pearson Product Moment Correlation' which is expressed as 'r' (Burns & Burns, 2008).
Since dependent and independent variables are expressed as scale data in this study, this coefficient is employed for measuring the relationship between them.

Varying between +1.00 and -1.00, coefficient 'r' expresses both the direction and the strength of a relationship. +1.00 indicates a perfect positive correlation and -1.00 a perfect negative correlation. The value 0 reveals that there is no relationship between variables (Burns & Burns, 2008).

Interpretation of the correlation size is controversial and no generally accepted schema exists. This research follows the schema presented by Burns & Burns (2008), which is presented below:

Table 6.18: Interpretation of correlation size (Burns & Burns, 2008, p. 346)

Correlation Coefficient (r)	Correlation	Description of relationship
± 0.90 – 1.00	Very high correlation	Very strong relationship
± 0.70 – 0.90	High correlation	Substantial relationship
± 0.40 – 0.70	Moderate correlation	Moderate relationship
± 0.20 – 0.40	Low correlation	Weak relationship
± 0.00 – 0.20	Slight correlation	Relationship so small as to be random

Besides investigating the correlation using Pearson's coefficient and the coefficient of determination, the variables need to be tested for their random correlation (Saunders et al., 2007). Testing the probability of a relationship is known as significance testing, expressed by the "p-value" or "p". Most researchers chose a significance level of 5.0% or even 1.0%.

In this research, the author decided to play it safe and applied a significance level of 1%. That means that the error probability of found relationships is smaller than 1%. Results of the correlation analyses and significance tests will be presented in a matrix showing the correlation coefficient of a dependent variable with the corresponding significance for each independent variable. The correlation coefficients conforming to the significance level are marked with two asterisks "**". Correlation coefficients which conform to the 5% level are marked with one asterisk "*". Any variables which do not match the 0.01 level will be regarded as not significant in this study.

6.3.1 RQ2-1: ANALYSIS & FINDINGS

RQ2-1 aims to identify the cost drivers influencing the required costs in the evaluation phase. Therefore, the internal and external personnel costs of a cost field during this lifecycle phase will be considered successively. Each cost field will be computed with a set of Pearson's correlations in order to determine if there are any significant relationships between the stated cost drivers.

INTERNAL PERSONNEL COSTS (Y1)

Correlation analysis identified 6 cost drivers with moderate or low relationships to the costs for internal personnel, which are “stability of organisation”, “number of total users”, “maturity of processes”, “number of locations” and the “satisfaction with ERP system”.

Table 6.19 gives an overview of the correlating cost drivers arranged in a descending order and sorted by their correlation coefficient from +1 to -1.

Table 6.19: Correlation between y1 and cost drivers during the evaluation phase

Independent Variables: Cost driver	**N**	**Pearson Corr.**	**r²**	**Signif. 2-tailed (p-Value)**	**Significance level**	**Interpretation**
x_{10} No. of total users	72	0.408	17%	0	**	moderate correlation
x_{1} No. of locations	72	0.344	12%	0.003	**	low correlation
x_{17} No. of external consultants	70	0.275	8%	0.021	*	low correlation
x_{12-8} Finance module	72	0.223	5%	0.059		low correlation
x_{13} Complexity data	72	0.172	3%	0.148		slight correlation
x_{12-4} PM module	72	0.164	3%	0.169		slight correlation
x_{8} No. of reports	67	0.143	2%	0.247		slight correlation
x_{16} No. of internal project members	70	0.141	2%	0.246		slight correlation
x_{4} Revenue	72	0.136	2%	0.264		slight correlation
x_{12-6} Sales module	72	0.133	2%	0.265		slight correlation
x_{5-2} ERP system - Microsoft	72	0.132	2%	0.268		slight correlation
x_{12-12} SCM module	72	0.129	2%	0.281		slight correlation
x_{12} No. of modules	72	0.126	2%	0.29		slight correlation
x_{12-14} PDA module	72	0.126	2%	0.29		slight correlation
x_{18} Ratio internal/external	69	0.106	1%	0.382		slight correlation
x_{12-2} MM module	72	0.082	1%	0.495		slight correlation
x_{5-3} ERP system - proALPHA	72	0.062	0%	0.605		no correlation
x_{12-11} Production module	72	0.054	0%	0.655		no correlation
x_{32} Complexity of bus. processes	72	0.050	0%	0.679		no correlation
x_{12-9} Accounting module	72	0.048	0%	0.692		no correlation
x_{23} Availability of management	69	0.044	0%	0.72		no correlation
x_{12-3} Calculation module	72	0.040	0%	0.741		no correlation
x_{12-5} DMS module	72	0.038	0%	0.753		no correlation
x_{11} No. of user groups	71	0.032	0%	0.791		no correlation
x_{9} No. of EDIs	70	0.009	0%	0.939		no correlation
x_{27} Critical attitude of users	70	-0.011	0%	0.93		no correlation
x_{5-1} ERP system - SAP	72	-0.011	0%	0.929		no correlation
x_{12-7} CRM module	72	-0.020	0%	0.866		no correlation
x_{14} Complexity of interfaces	70	-0.020	0%	0.868		no correlation

x_{22} Team composition	70	-0.030	0%	0.804		no correlation
x_{6} No. Of interfaces	69	-0.032	0%	0.795		no correlation
$x_{5\text{-}5}$ ERP system - ABAS	72	-0.033	0%	0.781		no correlation
x_{2} No. of org. units or depts.	72	-0.037	0%	0.758		no correlation
$x_{5\text{-}4}$ ERP system - SAGE	72	-0.039	0%	0.748		no correlation
x_{29} Employee involvement	70	-0.064	0%	0.597		no correlation
$x_{12\text{-}13}$ Detailed planning module	72	-0.065	0%	0.588		no correlation
x_{15} Complexity of reports	66	-0.069	0%	0.582		no correlation
x_{7} No. Of modifications	68	-0.070	0%	0.572		no correlation
$x_{12\text{-}10}$ HRM module	72	-0.079	1%	0.508		slight correlation
$x_{12\text{-}15}$ MDA module	72	-0.092	1%	0.442		slight correlation
$x_{5\text{-}6}$ ERP system - other	72	-0.092	1%	0.441		slight correlation
x_{20} Team quality	71	-0.094	1%	0.435		slight correlation
$x_{12\text{-}1}$ Purchasing module	72	-0.114	1%	0.341		slight correlation
x_{19} Fit system/organisation	71	-0.182	3%	0.129		slight correlation
x_{21} Team maturity	68	-0.200	4%	0.102		low correlation
x_{26} Consulting quality	68	-0.215	5%	0.078		low correlation
x_{28} User quality	68	-0.219	5%	0.072		low correlation
x_{24} Availability of business users	68	-0.232	5%	0.057		low correlation
x_{30} Management involvement	70	-0.234	5%	0.052		low correlation
x_{34} Willingness to change	70	-0.245	6%	0.041	*	low correlation
x_{25} Consulting experience	67	-0.282	8%	0.021	*	low correlation
x_{36} Satisfaction ERP system	71	-0.309	10%	0.009	**	low correlation
x_{31} Maturity of processes	69	-0.358	13%	0.003	**	low correlation
x_{35} Commitment management	71	-0.390	15%	0.001	**	low correlation
x_{33} Stability of organisation	70	-0.422	18%	0	**	moderate correlation

The table shows two moderate relationships:

The strongest one is “Stability of organisation” (x33) with a correlation coefficient of -0.422. This is significant at the 0.01 level, which means that this is no random connection but a real effect within the population. The stability within an organisation appears to provide a moderate guide to the internal personnel costs as it predicts 18% of the variance. The identified relationship indicates that the higher the stability of an organisation, the smaller the expense for internal personnel costs.

The second strongest relationship is the “number of total users” (x10), with a value of + 0.402. With a significance level of 0.01, the variable is very likely to have an effect within the population. With a coefficient of determination of 17%, the number of users explains the internal personal costs in the evaluation phase moderately. The finding means that the higher the number of total users, the higher the required expense for internal personnel costs. This variable has also been confirmed by previous studies.

The 4 remaining cost drivers “commitment management” (x35), “maturity of processes” (x31), “number of locations” (x1) and “satisfaction of an ERP system” (x36) feature a significance at the 0.01 level, but show only a low correlation.
With a value of -0.390, the independent variable “commitment management” (x35) is the strongest among the low correlation variables and is able to predict 15% of the variance of the required costs for internal personnel in the evaluation phase.

This value is followed by “maturity of processes” (x31) which features a correlation value of -0.358. Its coefficient of determination presents a value of 13%.

With a correlation coefficient of +0.344, the fifth strongest relationship was found in the variable “number of locations” (x1). Its coefficient of determination shows that the number of locations is able to explain 12% of the variance.

The lowest correlation coefficient was found for “satisfaction with ERP system” (x36). With a value of -0.309, this variable features a low negative relationship to the required internal personnel costs. With a coefficient of determination of 10%, “satisfaction with ERP system” explains the internal personnel costs during the evaluation phase only weakly.

All other cost drivers do not have any or only slight correlations.

So far, this chapter has identified 6 relevant relationships between the cost drivers and the internal personnel costs. They are summarised in table 6.20.

Table 6.20: Findings internal personnel costs during ERP evaluation phase (y1)

	Correlation (r)	r²	Correlation strength
Positive Correlations			
x10 No. of total users	+0.408	17%	moderate correlation
x1 No. of locations	+0.344	12%	low correlation
Negative Correlations			
x_{36} Satisfaction with ERP system	-0.309	10%	low correlation
x_{31} Maturity of processes	-0.358	13%	low correlation
x_{35} Commitment management	-0.390	15%	low correlation
x_{33} Stability of organisation	-0.422	18%	moderate correlation

Some results are rather expected ones, whilst others are very surprising:
Starting with the anticipated results, a positive correlation between the number of total users and costs for internal personnel is bound to exist, since the ERP requirements of each user need to be gathered, defined and analysed during the evaluation phase. Of course, this obviously involves internal personnel and to generate costs for internal personnel in consequence.
The positive relationship to the number of locations also seems very logical. As with the quantity of users, the requirements of each location need to be gathered, analysed and defined. This again occupies internal personnel and causes costs.

Regarding the negative relationships, the correlation between maturity of processes and commitment management seems to be logical.
The second chapter of this study describes the process-oriented, integrational feature of ERP systems. Having clear and defined processes appears to simplify the collection, analysis and definition of them and incurs fewer costs for internal personnel. However, the author of this study expected this influence to be higher.
Commitment management is about dedication, motivation and devotedness. These soft factors contributing to the success of an ERP project seems to be very likely. Apparently employees perform more effectively and more purposefully when they are committed to the company they work for. Nevertheless, also at this point, the author expected at least a moderate correlation for this variable.

The negative relationship to the variables "stability of organisation" and "satisfaction with ERP system", which means that the costs decrease with a higher stability and more satisfaction, was a very surprising finding.

One would have expected a positive relationship here; however, these results are a bit paradox and difficult to interpret. It would have been more logical that stable organisations are willing to spend more money on the evaluation process because they are moneyed and can afford to occupy their employees for selecting the best fitting system. Likewise, the variable "satisfaction with ERP System" would be more plausible in a positive connection. One should expect that the higher the expense for evaluation, the better the selection process and the better is the satisfaction with the ERP system.

So why would the costs for internal personnel decrease with a higher stability of the organisation? A possible answer could be found by regarding the correlation the other way round. Instable organisations spend more for internal personnel costs during the evaluation process. There might be several reasons for this: One reason might be that the processes are not very stable and no one really coordinates and manages the evaluation process within the company. This might lead to an ineffective performance. Another reason might be that such organisations hope to achieve stability by implementing an ERP system that fits their processes. Consequently, they are willing to spend more for the selection process.

Another paradox is the decreasing satisfaction with the ERP system by raising costs for internal personnel. As mentioned earlier, this would have been expected the other way round. A possible explanation could be the following: When companies spend a lot of money on the internal personnel during the evaluation phase, it could indicate that such companies aim to get a clear idea of their requirements, and that would involve many employees. This clear idea may result in a great disappointment if vendors deliver less than expected.

Another astonishing result is that some independent variables do not, or only slightly, influence the internal personnel costs during the evaluation phase. For example, the variables "number of modules", "number of internal project members" or "complexity of business processes" were expected to have a strong positive correlation. But, in fact, they only predict the variance about 2% and did not meet the significant level.

However, the correlation analysis is not able to provide a causal connection between variables. All possible explanations given are pure speculation. This analysis is more about finding suitable indicators which influence the costs of ERP systems. These indicators can, of course, be unapparent but it is a fact that they have a relation to the dependent variable.

Nevertheless, when regarding each variable separately, a large remainder of unexplained variance suggesting the existence of other factors that influence the relationship can be discovered. The question whether these factors are the other discovered cost drivers or completely different aspects cannot be answered at this point, since it requires another statistical approach, like multiple regression, which will be done later.
Regardless of the identified results, the evaluation phase seems to be the most difficult lifecycle phase for predicting the required costs. The extent of the costs depends on how much a company is willing to spend on selecting the ERP system that best matches its needs. Accordingly, it appears difficult to find a principle law which predicts the costs for this lifecycle phase.

EXTERNAL PERSONNEL COSTS (Y2)
A correlation analysis between the external personnel costs during the evaluation phase and the independent variables (Xi) discovered 7 relationships with moderate or low strength at the significance level 0.01. These are as follows: Commitment management, stability of organisation, satisfaction with ERP system, number of locations, number of users, maturity process, and team maturity.

Table 6.21 gives an overview of the correlating cost drivers arranged in descending order and sorted by their correlation coefficients from +1 to -1:

Table 6.21: Correlation between y2 and cost drivers during the evaluation phase

Independent Variables: Cost driver	**N**	**Pear-son Corr.**	**r²**	**Signif. 2-tailed (p-Value)**	**Signifi-cance Level**	**Interpretation**
x_1 No. of locations	72	0.406	16%	0	**	moderate correlation
x_{10} No. of total users	72	0.399	16%	0.001	**	low correlation
$x_{5\text{-}2}$ ERP system - Microsoft	72	0.242	6%	0.4	*	low correlation

x_{17} No. of external consultants	70	0.212	4%	0.78		low correlation
x_{4} Revenue	69	0.177	3%	0.145		slight correlation
x_{12-8} Finance module	72	0.172	3%	0.15		slight correlation
x_{12-12} SCM module	72	0.163	3%	0.171		slight correlation
x_{18} Ratio internal/external	69	0.156	2%	0.2		slight correlation
x_{13} Complexity of data	72	0.151	2%	0.205		slight correlation
x_{23} Availability of management	69	0.138	2%	0.257		slight correlation
x_{8} No. of reports	67	0.119	1%	0.338		slight correlation
x_{12-3} Calculation module	72	0.108	1%	0.367		slight correlation
x_{12-7} CRM module	72	0.089	1%	0.455		slight correlation
x_{12-14} PDA module	72	0.087	1%	0.465		slight correlation
x_{12-6} Sales module	72	0.084	1%	0.485		slight correlation
x_{12-11} Production module	72	0.074	1%	0.538		slight correlation
x_{11} No. of user groups	71	0.072	1%	0.549		slight correlation
x_{12} No. of modules	72	0.062	0%	0.604		no correlation
x_{12-2} MM module	72	0.038	0%	0.753		no correlation
x_{29} Employee involvement	70	0.026	0%	0.833		no correlation
x_{32} Complexity of bus. processes	72	0.018	0%	0.798		no correlation
x_{16} No. of internal project members	70	0.006	0%	0.96		no correlation
x_{14} Complexity of interfaces	70	0.004	0%	0.974		no correlation
x_{6} No. of interfaces	69	0.003	0%	0.98		no correlation
x_{12-1} Purchasing module	72	0.002	0%	0.988		no correlation
x_{2} No. of org. units or depts.	72	-0.003	0%	0.982		no correlation
x_{12-4} PM module	72	-0.004	0%	0.973		no correlation
x_{9} No. of EDIs	70	-0.014	0%	0.907		no correlation
x_{5-1} ERP system - SAP	72	-0.018	0%	0.878		no correlation
x_{7} No. of modifications	68	-0.022	0%	0.859		no correlation
x_{27} Critical attitude of users	70	-0.035	0%	0.774		no correlation
x_{5-4} ERP system - SAGE	72	-0.041	0%	0.731		no correlation
x_{12-5} DMS Module	72	-0.054	0%	0.652		no correlation
x_{5-3} ERP system - proALPHA	72	-0.058	0%	0.626		no correlation
x_{5-5} ERP system - ABAS	72	-0.063	0%	0.6		no correlation
x_{12-15} MDA module	72	-0.066	0%	0.585		no correlation
x_{12-9} Accounting module	72	-0.072	1%	0.55		slight correlation
x_{5-6} ERP system - other	72	-0.089	1%	0.459		slight correlation
x_{15} Complexity of reports	66	-0.118	1%	0.346		slight correlation
x_{12-10} HRM module	72	-0.123	2%	0.305		slight correlation
x_{12-13} Detailed planning module	72	-0.132	2%	0.269		slight correlation
x_{22} Team composition	70	-0.179	3%	0.139		slight correlation
x_{30} Management involvement	70	-0.223	5%	0.064		low correlation
x_{34} Willingness to change	70	-0.242	6%	0.044	*	low correlation
x_{20} Team quality	71	-0.242	6%	0.042	*	low correlation
x_{28} User quality	68	-0.244	6%	0.045		low correlation
x_{26} Consulting quality	68	-0.287	8%	0.018	*	low correlation

x_{24} Availability of business users	68	-0.289	8%	0.017	*	low correlation
x_{19} Fit system/organisation	71	-0.296	9%	0.012	*	low correlation
x_{25} Consulting experience	67	-0.303	9%	0.013	*	low correlation
x_{21} Team maturity	68	-0.346	12%	0.004	**	low correlation
x_{31} Maturity processes	69	-0.361	13%	0.002	**	low correlation
x_{36} Satisfaction with ERP system	71	-0.426	18%	0	**	moderate correlation
x_{33} Stability of organisation	70	-0.455	21%	0	**	moderate correlation
x_{35} Commitment management	71	-0.46	21%	0	**	moderate correlation

The results presented in this table are similar to those of the analysis of internal personnel costs during the evaluation phase (y1) in the chapter before: Rather the same variables were identified to have a correlation and their strength is rather similar to the ones discovered in the chapter before. One "new" variable was discovered to have a correlation to external personnel costs, which is the variable "team maturity".

In contrast to the previous chapter, 2 more variables with moderate connection were found, which means there are 4 in total.
With a value of +0.460, the strongest correlation is the relationship to "commitment management" (x35). This is a guideline for the internal personnel costs as it explains 21% of the variance. This variable only showed a low correlation in the previous chapter.

This connection is closely followed by the second strongest relationship to "stability of organisation". It has a correlation coefficient of -0.455 and also explains 21% of the variance.
Having a correlation coefficient of -0.426, "satisfaction with ERP system" is the third strongest connection. With a coefficient of determination of 18%, satisfaction with the ERP system moderately explains the internal personnel costs during the evaluation phase.
The fourth variable with a moderate relationship is "number of locations". Having a correlation coefficient of 0.40, its coefficient of determination predicts 16% of the variance.

The remaining 3 variables featuring a low correlation are the variables "number of users", "maturity of processes" and "team maturity".
These are headed by "number of users" (x33) with a correlation coefficient of +0.399. The number of users within an organisation seems to provide a low to moderate guide to the external personnel costs as it predicts 16% of the variance, in any case.

This is followed by two negative relationships, the “maturity of processes” and “team maturity”. “Maturity of processes” has a correlation coefficient of -0.361 and coefficient of determination of 13%. The lowest relationship at the significance level 0.01 is ascertained for “team maturity”. It is -0.346. The relationship is low with a coefficient of determination of 12%.

At the significance level of 0.01, all other variables show only a slight or no correlation to the external personnel costs during the evaluation phase.

Significant relationships at a significance level of 0.01 could be identified in this section 6. The table below summarises them:

Table 6.22: Findings external personnel costs during ERP evaluation phase (y2)

	Correlation (r)	r^2	Correlation strength
Positive correlations			
x1 No. of locations	0.406	16%	moderate correlation
x10 No. of total users	0.399	16%	low correlation
Negative Correlations			
x_{21} Team maturity	-0.346	12%	low correlation
x_{31} Maturity of processes	-0.361	13%	low correlation
x_{36} Satisfaction with ERP system	-0.426	18%	moderate correlation
x_{33} Stability of organisation	-0.455	21%	moderate correlation
x_{35} Commitment management	-0.46	21%	moderate correlation

As mentioned earlier, the correlation result of external personnel costs shows strong similarities to internal personnel costs. In consequence, the interpretation of results is also related. As with the internal personnel costs, the results are likely for some variables and astonishing for others:

As with internal personnel costs, a positive correlation between the number of locations and costs for external personnel seems to be very obvious. Since the requirements of each location need to be compiled, analysed and defined, the costs for external personnel increases with the number of each location when engaging external personnel for consulting.

As with the number of locations, the positive connection between the “number of total users” and costs for external personnel is expected, because the ERP requirements of each user need to be compiled, defined and analysed during the evaluation phase. Retraining external personnel for doing these tasks generates costs.

Aside from the team maturity, all negative correlations have been discussed in the previous section and can be transferred to the external personnel costs, since the task is the same, regardless of whether it is conducted by internal or external personnel.
The mystery is the negative relationship to the indicators “stability of organisation” and “satisfaction with ERP system” which means that the costs decrease with a higher stability and more satisfaction remains. Possible explanations were given in the previous section and can be transferred to this one.

The identified negative relationship to team maturity appears to be logical. Having a team of experts who are familiar with the processes and products and who are willing to work effectively and purposefully with the external personnel decreases the costs for external staff during the evaluation phase.
It is astonishing that this variable is not relevant for the amount of internal personnel costs during the evaluation phase.

To conclude this section, this interpretation should not hide the fact that each variable has a high remainder of unexplained variance. As with internal personnel costs, it remains unanswered if this variance can be explained by the other identified cost drivers or if completely different factors are needed.
It is a fact that external personnel costs have a dependence relationship to the 7 stated cost drivers.

6.3.2 RQ2-2: ANALYSIS & FINDINGS

In this sub-section, the relevant cost drivers during the implementation phase will be identified. To this end, each cost field during the implementation phase – which are internal personnel costs, external personnel costs, ERP software costs, licence costs, and hardware costs – will be analysed by way of correlation analysis. This chapter will present the cost fields one after another.

INTERNAL PERSONNEL COSTS (Y3)

The correlation analysis identified a relationship between 9 cost drivers and the internal personnel costs during the implementation phase (y3).

The outcome of the analysis is presented in table 6.23, arranged in a descending order sorted by their correlation coefficients from +1 to -1.

Table 6.23: Correlation between y3 and cost drivers during implementation phase

Independent Variable: Cost Driver	N	Pearson Corr.	r^2	Signif. 2-tailed (p-Value)	Signifi-cance level	Interpretation
x_{10} No. of total users	72	0.693	48%	0	**	moderate correlation
x_{17} No. of external consultants	70	0.643	41%	0	**	moderate correlation
x_{16} No. of internal project members	70	0.526	28%	0	**	moderate correlation
x_2 No. of org. units or depts.	72	0.500	25%	0	**	moderate correlation
x_{5-1} ERP system - SAP	72	0.465	22%	0	**	moderate correlation
x_4 Revenue	69	0.391	15%	0.001	**	low correlation
x_1 No. of locations	72	0.368	14%	0.001	**	low correlation
x_{12-5} DMS module	72	0.243	6%	0.04	*	low correlation
x_8 No. of reports	67	0.234	5%	0.057		low correlation
x_{12} No. of modules	72	0.230	5%	0.051		low correlation
x_{13} Complexity of data	72	0.206	4%	0.082		low correlation
x_{12-12} SCM module	72	0.190	4%	0.11		slight correlation
x_{12-8} Finance module	72	0.183	3%	0.124		slight correlation
x_{12-9} Accounting module	72	0.181	3%	0.128		slight correlation
x_{12-4} PM module	72	0.179	3%	0.133		slight correlation
x_{12-11} Production module	72	0.166	3%	0.163		slight correlation
x_{12-3} Calculation module	72	0.149	2%	0.211		slight correlation
x_{12-6} Sales module	72	0.113	1%	0.344		slight correlation
x_{21} Team maturity	68	0.112	1%	0.365		slight correlation
x_9 No. of EDIs	70	0.110	1%	0.365		slight correlation
x_{5-3} ERP system - proALPHA	72	0.103	1%	0.39		slight correlation
x_6 No. of interfaces	69	0.095	1%	0.435		slight correlation
x_{12-10} HRM module	72	0.091	1%	0.448		slight correlation
x_{12-1} Purchasing module	72	0.089	1%	0.459		slight correlation
x_{12-2} MM module	72	0.084	1%	0.485		slight correlation
x_{14} Complexity of interfaces	72	0.068	0%	0.578		no correlation
x_{20} Team quality	71	0.042	0%	0.729		no correlation
x_{5-2} ERP system - Microsoft	72	0.027	0%	0.825		no correlation
x_{22} Team composition	70	0.024	0%	0.846		no correlation
x_{15} Complexity of reports	66	0.023	0%	0.852		no correlation
x_{28} User quality	68	0.014	0%	0.907		no correlation
x_{11} No. of user groups	71	0.002	0%	0.989		no correlation
x_{32} Complexity of bus. processes	72	-0.001	0%	0.79		no correlation
x_{29} Employee involvement	70	-0.031	0%	0.8		no correlation
x_{12-14} PDA module	72	-0.032	0%	0.789		no correlation
x_{5-5} ERP system - ABAS	72	-0.038	0%	0.753		no correlation
x_{19} Fit of system/organization	71	-0.048	0%	0.689		no correlation
x_{18} Ratio internal/external	69	-0.050	0%	0.679		no correlation
x_7 No. of modifications	68	-0.056	0%	0.649		no correlation

x_{26} Consulting quality	68	-0.065	0%	0.596		no correlation
x_{25} Consulting experience	67	-0.075	1%	0.548		slight correlation
x_{12-7} CRM module	72	-0.081	1%	0.497		slight correlation
x_{27} Critical attitude of users	70	-0.083	1%	0.495		slight correlation
x_{23} Availability of management	69	-0.092	1%	0.452		slight correlation
x_{12-15} MDA module	72	-0.098	1%	0.414		slight correlation
x_{34} Willingness to change	70	-0.104	1%	0.392		slight correlation
x_{30} Management involvement	70	-0.113	1%	0.351		slight correlation
x_{36} Satisfaction with ERP system	71	-0.116	1%	0.335		slight correlation
x_{5-4} ERP system - SAGE	72	-0.119	1%	0.319		slight correlation
x_{12-13} Detailed planning module	72	-0.130	2%	0.276		slight correlation
x_{24} Availability of business users	68	-0.214	5%	0.08		low correlation
x_{31} Maturity of processes	69	-0.272	7%	0.024	*	low correlation
x_{35} Commitment management	71	-0.299	9%	0.011	*	low correlation
x_{5-6} ERP system - other	72	-0.334	11%	0.004	**	low correlation
x_{33} Stability of organisation	70	-0.345	12%	0.003	**	low correlation

The table shows 5 moderate and 4 low relationships at the 0.01 significance level:
The variable “total number of users” shows the strongest relationship with a correlation coefficient of 0.693. With a coefficient of determination of 48%, the number of users seems to make up a huge portion of the internal personnel costs during the evaluation phase.

This variable is followed by the number of external consultants. The correlation between this number and the internal personnel costs is 0.643 and, surprisingly, provides a moderate guide to this cost field. Its coefficient of determination predicts 41% of the variance, which is determined as a moderate connection.

The third highest relationship can be found at the variable “number of internal project members” (x16). This variable presents a significant correlation factored with 0.526. The number of internal project members within an organisation also seems to provide a moderate guide to the internal personnel costs as it predicts 28% of the variance.

This variable is followed by the “number of departments or units” which shows a correlation of 0.500 and explains the variance of the costs for internal personnel during the implementation process with 25%.

With a correlation of 0.465, the last moderate relationship is identified between “ERP System – SAP” and y3. This indicates that organisations choosing SAP have higher costs for internal personnel during the implementation phase, especially because the variable “ERP-others” shows a low, but significant negative relationship, which implies that other vendors (beside the famous ones, like Microsoft, proALPHA, SAGE, and ABAS, who receive their own sub-variable) cause lower costs for internal personnel at this stage of the lifecycle.

The 4 variables with a low correlation are namely “revenue” (x4), “number of locations” (x1), “stability of organisation” (x33) and “ERP system - other”.

As before, the analysis found some foreseeable results but also provided some unexpected outcomes. Table 6.24 below summarises the findings:

Table 6.24: Findings internal personnel costs during ERP implement. phase (y3)

	Correlation (r)	r^2	Correlation Strength
Positive Correlations			
x10 No. of total users	0.693	48%	moderate correlation
x17 No. of external consultants	0.643	41%	moderate correlation
x16 No. of internal project members	0.526	28%	moderate correlation
x2 No. of org. units or depts.	0.500	25%	moderate correlation
x5-1 ERP system - SAP	0.465	22%	moderate correlation
x4 Revenue	0.391	15%	low correlation
x1 No. of locations	0.368	14%	low correlation
Negative Correlations			
$x_{5\text{-}6}$ ERP system - other	-0.334	11%	low correlation
x_{33} Stability of organisation	-0.345	12%	low correlation

The positive correlations between “number of total users”, “number of internal project members”, “number of organisational units or departments” and “number of locations” and the internal personnel costs during the implementation phase were quite expected. Since every user is more or less involved in the project and needs to be trained during this phase, it appears logical that costs for internal personnel increase with the number of users. The same argument applies to the “number of internal project members”. The more people are involved in the project, the higher the costs for internal personnel.

Both variables, the “number of organisational units or departments” and “number of locations”, might be explained by the rule: the more needs to be implemented, the higher the costs. The number of organisational departments or units might focus on implementing a variety of things whilst the number of locations indicates the quantity. Nevertheless, the variable “number of modules” does not play a significant role in this cost field.

The low positive relationship with the variable “revenue” is logical when understanding it as an indicator of the project size. In general, one assumes that the costs increase with the size of the project.

Regarding the surprising results, one of the most surprising is the combination of the positive correlation “ERP System – SAP” and the negative relationship with “ERP-others” to the internal personnel costs. As mentioned above, this indicates that SAP causes higher costs whilst other vendors (beside the famous ones, like Microsoft, proALPHA, SAGE, and ABAS, which receive their own sub-variable) cause lower costs for internal personnel at this stage of the lifecycle. Explaining this would require analysing the different approaches of large vendors in contrast to those of the smaller ones, which cannot be done in this thesis.

Another very astonishing result is the relatively strong positive connection to “no. of external consultants”. At first glance, this seems to be rather contradictory and inconsistent. Why should internal personnel costs increase with the number of external staff? One would expect it to be the other way round. A possible explanation could be that consultants rely on the cooperation and feedback of internal staff. The more external staff are engaged, the more cooperation is required, which occupies the internal staff.

The astonishing result of the variable “stability of organisation” appears again at this stage. Possible explanations are given in RQ1-1 which can be transferred to this section.

EXTERNAL PERSONNEL COSTS (Y4)

The dependent variable “external personnel costs during the implementation phase” shows a significant connection and a positive relationship to 11 cost drivers, as follows: “number of external consultants”, “number of total users”, “number of internal project managers", "ERP System – SAP", "number of organisations or departments", "accounting module", "Human Resource Management module", "revenue", "number of locations". A negative correlation is identified between "stability of organisation" as well as "ERP system – other" and y4. These relationships are significant at the 0.01 level.

Table 6.25 shows an overview of the correlating cost drivers arranged in a descending order sorted by their correlation coefficient from +1 to -1:

Table 6.25: Correlation between y4 and cost drivers during implementation phase

Independent Variable: Cost Drivers	N	Pearson Corr.	r^2	Signif. 2-tailed (p-Value)	Significance level	Interpretation
x_{17} No. of external consultants	70	0.719	52%	0	**	substantial correlation
x_{10} No. of total users	72	0.612	37%	0	**	moderate correlation
x_{16} No. of internal project members	70	0.567	32%	0	**	moderate correlation
x_{5-1} ERP system - SAP	72	0.482	23%	0	**	moderate correlation
x_2 No. of org. units or depts.	72	0.443	20%	0	**	moderate correlation
x_{12-9} Accounting module	72	0.383	15%	0.001	**	low correlation
x_{12-10} HRM module	72	0.374	14%	0.001	**	low correlation
x_4 Revenue	69	0.365	13%	0.002	**	low correlation
x_1 No. of locations	72	0.356	13%	0.002	**	low correlation
x_{12} No. of modules	72	0.291	8%	0.013	*	low correlation
x_{12-8} Finance module	72	0.288	8%	0.014	*	low correlation
x_{5-2} ERP system - Microsoft	72	0.263	7%	0.026	*	low correlation
x_8 No. of reports	67	0.257	7%	0.035	*	low correlation
x_{13} Complexity of data	72	0.240	6%	0.042	*	low correlation
x_{12-11} Production module	72	0.194	4%	0.103		slight correlation
x_{12-5} DMS module	72	0.190	4%	0.109		slight correlation
x_{12-12} SCM module	72	0.181	3%	0.128		slight correlation
x_{12-3} Calculation module	72	0.166	3%	0.163		slight correlation
x_6 No. of interfaces	69	0.156	2%	0.201		slight correlation
x_9 No. of EDIs	70	0.153	2%	0.206		slight correlation
x_{12-6} Sales module	72	0.136	2%	0.255		slight correlation
x_{32} Complexity of bus. processes	72	0.112	1%	0.349		slight correlation
x_{12-1} Purchasing module	72	0.094	1%	0.433		slight correlation
x_{12-2} MM module	72	0.091	1%	0.448		slight correlation
x_{14} Complexity of interfaces	70	0.053	0%	0.663		no correlation
x_{18} Ratio internal/external	69	0.050	0%	0.69		no correlation
x_{22} Team composition	70	0.050	0%	0.682		no correlation
x_{12-4} PM module	72	0.035	0%	0.771		no correlation
x_{26} Consulting quality	68	0.029	0%	0.815		no correlation
x_{5-3} ERP system - proALPHA	72	0.020	0%	0.867		no correlation
x_{11} No. of user groups	71	0.011	0%	0.927		no correlation
x_{15} Complexity of reports	66	-0.005	0%	0.97		no correlation
x_7 No. of modifications	68	-0.032	0%	0.799		no correlation
x_{29} Employee involvement	70	-0.044	0%	0.715		no correlation

x_{24} Availability of business users	68	-0.045	0%	0.716		no correlation
x_{21} Team maturity	68	-0.047	0%	0.702		no correlation
x_{30} Management involvement	70	-0.050	0%	0.679		no correlation
x_{23} Availability of management	69	-0.070	0%	0.57		no correlation
$x_{5\text{-}5}$ ERP system - ABAS	72	-0.074	1%	0.534		slight correlation
$x_{12\text{-}15}$ Modul MDA	72	-0.077	1%	0.518		slight correlation
x_{36} Satisfaction ERP system	71	-0.077	1%	0.522		slight correlation
x_{20} Team quality	71	-0.078	1%	0.518		slight correlation
$x_{12\text{-}7}$ CRM module	72	-0.079	1%	0.509		slight correlation
x_{25} Consulting experience	67	-0.083	1%	0.502		slight correlation
x_{34} Willingness to change	70	-0.097	1%	0.423		slight correlation
$x_{12\text{-}13}$ Detailed planning module	72	-0.118	1%	0.322		slight correlation
x_{28} User quality	68	-0.141	2%	0.25		slight correlation
x_{19} Fit system/organisation	71	-0.165	3%	0.17		slight correlation
$x_{12\text{-}14}$ PDA module	72	-0.170	3%	0.153		slight correlation
$x_{5\text{-}4}$ ERP system - SAGE	72	-0.171	3%	0.151		slight correlation
x_{27} Critical attitude of users	70	-0.179	3%	0.138		slight correlation
x_{35} Commitment management	71	-0.215	5%	0.071		low correlation
x_{31} Maturity of processes	69	-0.296	9%	0.014	*	low correlation
x_{33} Stability of organisation	70	-0.307	9%	0.01	**	low correlation
$x_{5\text{-}6}$ ERP system - other	72	-0.438	19%	0	**	moderate correlation

The table shows 1 substantial, 5 moderate and 4 low relationships.
The substantial connection is identified between the “number of external consultants” and external personnel costs. The correlation between them is 0.719. It seems to provide a high guideline for the external personnel costs as it predicts 52% of the variance.

The strongest moderate connection is discovered for “number of total user”. With a correlation coefficient of 0.612, its coefficient of determination explains 37% of the variance. This is followed by the “number of internal project managers”, “ERP System – SAP”, “number of organisations or departments” and the negative correlation “ERP system – others”.

The following cost drivers show low positive relations to y4:
“Accounting module”, “Human Resource Management module”, “revenue”, and “number of locations”. The variable “stability of organisation” shows a low negative correlation again.

So far, this chapter has discovered 11 cost drivers which have a relevant connection to the external personnel costs during the implementation phase.
They are summarised in the following table:

Table 6.26: Findings external personnel costs during ERP implement. phase (y4)

	Correlation (r)	r^2	Correlation Strength
Positive Correlation			
x17 No. of external consultants	0.719	52%	substantial correlation
x_{10} No. of total users	0.612	37%	moderate correlation
x_{16} No. of internal project members	0.567	32%	moderate correlation
x_{5-1} ERP system - SAP	0.482	23%	moderate correlation
x_2 No. of org. units or depts.	0.443	20%	moderate correlation
x_{12-9} Accounting module	0.383	15%	low correlation
x_{12-10} HRM module	0.374	14%	low correlation
x_4 Revenue	0.365	13%	low correlation
x_1 No. of locations	0.356	13%	low correlation
Negative Correlation			
x33 Stability of organisation	-0.307	9%	low correlation
x5-6 ERP system - other	-0.438	19%	moderate correlation

Unsurprisingly, the variable "number of external consultants" shows the highest correlation. This seems to be very logical and does not need any further explanation.

The identified connections between y4 and the modules for accounting and human resource management are very interesting. This connection indicated that companies implementing these modules would expect a cost increase of external personnel costs. This could be explained by the wide difference of processes between companies, especially concerning these financial topics that might require more tailoring of the modules.

Again, the correlation analysis found a relationship between higher costs for y4 and SAP and lower expenses when choosing one of the smaller vendors, who did not get their own variable. This finding is described in the previous chapter and can be transferred to this finding. The variable "revenue" is also repeated and can be interpreted as above.

The remaining two variables are the surprising ones:
The correlation between the "number of internal project members" and y4 does not seem to make sense. Nevertheless, this study found the same connection within the internal personnel costs during the implementation phase, so obviously one variable induces the

other. One explanation might be that the higher the number of internal members, the higher their capacity to occupy external staff, which increases the costs.

Yet the analysis found a negative correlation to the variable “stability of organisation” which means that the costs decrease with a higher stability. Possible explanations were given in RQ2.1 and can be transferred to this one.

ERP SOFTWARE COSTS (Y5)

The correlation analysis for y5 identified 8 cost drivers at the significance level 0.01, which are “stability of organisation”, “number of total users”, “maturity of processes”, “number of locations” and “satisfaction with ERP system”.

Table 6.27 gives an overview of the correlating cost drivers arranged in a descending order sorted by their correlation coefficient from +1 to -1.

Table 6.27: Correlation between y5 and cost drivers during implementation phase

Independent Variables: Cost Drivers	**N**	**Pearson Corr.**	**r^2**	**Signif. 2-tailed (p-Value)**	**Significance level**	**Interpretation**
x_4 Revenue	69	0.543	29%	0	**	moderate correlation
x_{10} No. of total users	72	0.536	29%	0	**	moderate correlation
x_{17} No. of external consultants	70	0.383	15%	0.001	**	low correlation
x_2 No. of org. units or depts.	72	0.346	12%	0.003	**	low correlation
x_1 No. of locations	72	0.342	12%	0.003	**	low correlation
x_{11} No. of user groups	71	0.322	10%	0.006	**	low correlation
x_{16} No. of internal project members	70	0.322	10%	0.006	**	low correlation
x_{12-8} Finance module	72	0.294	9%	0.012	*	slight correlation
x_{12-5} DMS module	72	0.243	6%	0.039	*	slight correlation
x_{12-12} SCM module	72	0.243	6%	0.04	*	slight correlation
x_{13} Complexity of data	72	0.233	5%	0.049	*	slight correlation
x_{12} No. of modules	72	0.222	5%	0.061		slight correlation
x_{5-1} ERP system - SAP	72	0.177	3%	0.136		slight correlation
x_{12-10} HRM module	72	0.168	3%	0.159		slight correlation
x_{18} Ratio internal/external	69	0.163	3%	0.18		slight correlation
x_{12-4} PM module	72	0.142	2%	0.233		slight correlation
x_{12-9} Accounting module	72	0.138	2%	0.246		slight correlation
x_{32} Complexity of bus. processes	72	0.119	1%	0.957		slight correlation

$x_{12\text{-}3}$ Calculation module	72	0.113	1%	0.343		slight correlation
$x_{5\text{-}2}$ ERP system - Microsoft	72	0.106	1%	0.377		slight correlation
x_8 No. of reports	67	0.102	1%	0.413		slight correlation
$x_{12\text{-}11}$ Production module	72	0.098	1%	0.415		slight correlation
x_{27} Critical attitude of users	70	0.095	1%	0.434		slight correlation
$x_{12\text{-}2}$ MM module	72	0.087	1%	0.468		slight correlation
$x_{5\text{-}5}$ ERP system - ABAS	72	0.083	1%	0.489		slight correlation
$x_{5\text{-}3}$ ERP system - proALPHA	72	0.031	0%	0.793		no correlation
x_6 No. of interfaces	69	0.029	0%	0.812		no correlation
x_{29} Employee involvement	70	0.021	0%	0.86		no correlation
x_{15} Complexity of reports	66	0.018	0%	0.887		no correlation
x_9 No. of EDIs	70	0.011	0%	0.928		no correlation
$x_{12\text{-}14}$ PDA module	72	0.002	0%	0.984		no correlation
x_{24} Availability of business users	68	-0.002	0%	0.984		no correlation
x_{23} Availability of management	69	-0.008	0%	0.948		no correlation
$x_{12\text{-}6}$ Sales module	72	-0.014	0%	0.908		no correlation
x_{28} User quality	68	-0.024	0%	0.847		no correlation
$x_{12\text{-}7}$ CRM module	72	-0.028	0%	0.813		no correlation
x_7 No. of modifications	68	-0.035	0%	0.776		no correlation
$x_{12\text{-}13}$ Detailed planning module	72	-0.035	0%	0.773		no correlation
x_{14} Complexity of interfaces	70	-0.036	0%	0.77		no correlation
x_{34} Willingness to change	70	-0.038	0%	0.753		no correlation
x_{19} Fit of system/organisation	71	-0.058	0%	0.63		no correlation
x_{22} Team composition	70	-0.080	1%	0.509		slight correlation
$x_{5\text{-}4}$ ERP system - SAGE	72	-0.081	1%	0.5		slight correlation
x_{21} Team maturity	68	-0.109	1%	0.378		slight correlation
x_{31} Maturity of processes	69	-0.114	1%	0.003		slight correlation
x_{30} Management involvement	70	-0.129	2%	0.287		slight correlation
x_{36} Satisfaction with ERP system	71	-0.130	2%	0.279		slight correlation
x_{20} Team quality	71	-0.138	2%	0.25		slight correlation
$x_{12\text{-}1}$ Purchasing module	72	-0.147	2%	0.216		slight correlation
x_{25} Consulting experience	67	-0.183	3%	0.139		slight correlation
x_{35} Commitment management	70	-0.214	5%	0.073		low correlation
x_{26} Consulting quality	68	-0.220	5%	0.072		low correlation
$x_{5\text{-}6}$ ERP system - other	72	-0.229	5%	0.053		low correlation
$x_{12\text{-}15}$ MDA module	72	-0.245	6%	0.038	*	low correlation
x_{33} Stability of organisation	70	-0.389	15%	0.001	**	low correlation

The analysis found just two moderate and six low relationships.

The strongest relationship is measured between the variable “revenue” (x4). The correlation between the turnover and the expenditure for ERP software is 0.543. With a coefficient of determination of 29%, this variable delivers a moderate guide for assessing the cost field y5. This value is closely followed by the “number of total users”, which features also a coefficient of determination of 29% but has a slightly lower correlation coefficient of 0.536.

The following six variables present at least a low, but significant relationship:
Beginning with the positive correlation, this study found five, which are: number of external consultants, number of units or departments, number of locations, number of user groups, number of internal project members. The only negative correlation is the stability of organisation.
The strongest among them is the variable “external consultants” which correlates with the value 0.383 and shows a coefficient of determination of 15%. This result is very surprising since there seems to be no obvious dependence between costs for ERP software and external consultants. The other unexpected connection is the one to the number of internal project members. It correlates with a value of 0.322 and predicts 10% of the variance. The number of internal project members seems to be quite a good indicator of the “project size” for ERP project costs during the implementation phase.

The third surprising relationship is the negative connection to stability of organisation. With a coefficient of determination of 15%, it remains a mystery why the software expenses increase with the stability of the organisation.

The other variables, on the other hand, appear to be very logical. The correlation between the number of units or departments and y5 is 0.346 and predicts 12% of the variance. It seems that the higher the number of departments, the more software is needed and the higher the expense. This argument can also be applied to the variable “number of locations”, which correlates with the value 0.34 and also predicts 12%, as well as to the variable “number of user groups” which features a correlation coefficient of 0.322 and a coefficient of determination of 10%.

To conclude this section, the findings are summarised in the table below:

Table 6.28: Findings ERP software costs during ERP implementation phase (y5)

	Correlation (r)	r^2	Correlation Strength
Positive Correlation			
x_4 Revenue	0.543	29%	moderate correlation
x_{10} No. of total users	0.536	29%	moderate correlation
x_{17} No. of external consultants	0.383	15%	low correlation
x_2 No. of org. units or depts.	0.346	12%	low correlation
x_1 No. of locations	0.342	12%	low correlation
x_{11} No. of user groups	0.322	10%	low correlation
x_{16} No. of internal project members	0.322	10%	low correlation
Negative Correlation			
x33 Stability of organisation	-0.389	15%	low correlation

One has to admit that for each variable there is a high remainder of unexplained variance. It cannot be answered if this variance is explained by the other identified cost drivers or if completely other factors are needed. It is a fact that software costs are dependent upon the 8 stated cost drivers.

LICENCE COSTS (Y6)

Eight variables were identified to have a significant relationship with the licence costs (y6) at the significance level 0.01. The ones with positive correlations are: number of external consultants, number of locations, number of internal project managers, revenue, total number of users, number of units or departments, and the finance module. The only negative relationship occurs with the variable "maturity of processes".

Five of them hold a moderate correlation, whereas three have at least a low correlation. The results are presented in table 6.29.

Table 6.29: Correlation between y6 and cost drivers during implementation phase

Independent Variables: Cost Drivers	N	Pearson Corr.	r^2	Signif. 2-tailed (p-Value)	Significance level	Interpretation
x_{17} No. of external consultants	70	0.519	27%	0	**	moderate correlation
x_1 No. of locations	72	0.479	23%	0	**	moderate correlation
x_{16} No. of internal project members	70	0.477	23%	0	**	moderate correlation
x_4 Revenue	69	0.464	22%	0	**	moderate correlation

x_{10} No. of total users	72	0.440	19%	0	**	moderate correlation
x_{2} No. of org. units or depts.	72	0.314	10%	0.007	**	low correlation
x_{12-8} Finance module	72	0.313	10%	0.0008	**	low correlation
x_{8} No. of reports	67	0.260	7%	0.033	*	low correlation
x_{11} No. of user groups	71	0.258	7%	0.03	*	low correlation
x_{12-12} SCM module	72	0.243	6%	0.04	*	low correlation
x_{5-3} ERP system - proALPHA	72	0.227	5%	0.055		low correlation
x_{5-1} ERP system - SAP	72	0.206	4%	0.082		low correlation
x_{13} Complexity data	72	0.161	3%	0.176		slight correlation
x_{5-2} ERP system - Microsoft	72	0.131	2%	0.272		slight correlation
x_{21} Team maturity	68	0.119	1%	0.332		slight correlation
x_{20} Team quality	71	0.106	1%	0.38		slight correlation
x_{18} Ratio internal/external	69	0.092	1%	0.45		slight correlation
x_{12-5} DMS module	72	0.083	1%	0.486		slight correlation
x_{12-2} MM module	72	0.081	1%	0.499		slight correlation
x_{12} No modules	72	0.062	0%	0.603		no correlation
x_{22} Team composition	70	0.056	0%	0.646		no correlation
x_{6} No. of interfaces	69	0.050	0%	0.684		no correlation
x_{12-4} PM module	72	0.048	0%	0.686		no correlation
x_{12-9} Accounting module	72	0.024	0%	0.842		no correlation
x_{9} No. of EDIs	70	0.019	0%	0.875		no correlation
x_{12-11} Production module	72	0.018	0%	0.883		no correlation
x_{14} Complexity of interfaces	70	0.012	0%	0.925		no correlation
x_{12-15} MDA module	72	0.011	0%	0.924		no correlation
x_{7} No. of modifications	68	0.006	0%	0.961		no correlation
x_{12-10} HRM module	72	0.006	0%	0.96		no correlation
x_{15} Complexity of reports	66	-0.018	0%	0.889		no correlation
x_{28} User quality	68	-0.045	0%	0.713		no correlation
x_{12-6} Sales module	72	-0.048	0%	0.69		no correlation
x_{30} Management involvement	70	-0.055	0%	0.651		no correlation
x_{24} Availability of business users	68	-0.056	0%	0.65		no correlation
x_{27} Critical attitude of users	70	-0.056	0%	0.644		no correlation
x_{12-14} PDA module	72	-0.071	1%	0.553		slight correlation
x_{19} Fit of system/organisation	71	-0.082	1%	0.498		slight correlation
x_{12-3} Calculation module	72	-0.083	1%	0.487		slight correlation
x_{12-13} Detailed planning module	72	-0.096	1%	0.424		slight correlation
x_{23} Availability of management	69	-0.102	1%	0.406		slight correlation
x_{12-1} Purchasing module	72	-0.103	1%	0.388		slight correlation
x_{36} Satisfaction with ERP system	71	-0.107	1%	0.373		slight correlation
x_{5-5} ERP system - ABAS	72	-0.108	1%	0.365		slight correlation
x_{34} Willingness to change	70	-0.124	2%	0.307		slight correlation
x_{32} Complexity of bus. processes	72	-0.125	2%	0.053		slight correlation
x_{29} Employee involvement	70	-0.157	2%	0.194		slight correlation

$x_{5\text{-}4}$ ERP system - SAGE	72	-0.159	3%	0.182		slight correlation
$x_{12\text{-}7}$ CRM module	72	-0.165	3%	0.165		slight correlation
x_{25} Consulting experience	67	-0.171	3%	0.167		slight correlation
x_{35} Commitment management	71	-0.180	3%	0.133		slight correlation
x_{26} Consulting quality	68	-0.181	3%	0.141		slight correlation
x_{33} Stability of organisation	70	-0.181	3%	0.133		slight correlation
$x_{5\text{-}6}$ ERP system - other	72	-0.240	6%	0.042	*	moderate correlation
x_{31} Maturity of processes	69	-0.347	12%	0.003	**	moderate correlation

The results presented in the table are similar to those of the analysis of the software costs during the implementation phase (y5) in the previous chapter: rather the same variables were identified to have a connection. Just the “number of user groups” and “stability of organisation” are replaced with the modules “finance” and “team maturity”.

This result is very surprising since one expects the variables “number of users” and “number of modules” to be the ones with the highest correlation coefficient. These variables usually represent the main factors from which licence costs emerge.
However, the number of users appears to have just a moderate connection. The correlation between it and y6 is just 0.440, and the coefficient of determination can only predict 19% of the variance. The variable “number of modules” does not play a significant role at all. That seems to be a very contradictory and inconsistent result.

In contrast to the chapter before, three more variables with moderate connection could be found, which means there are five in total.
The four strongest ones are difficult to explain since there is no obvious relation between those variables and the licence costs:
The strongest correlation with a value of +0.519 is the relationship to “external personnel”. This provides a moderate guide to the licence costs as it predicts 27% of the variance. This result is very surprising since there appears to be no obvious connection between costs for licence and the number of external staff. An interpretation is difficult and an explanation for that would be pure speculation.
The same confusion occurs when regarding the second strongest correlation, which is found between the number of locations and y6. Having a correlation coefficient of 0.479 and a coefficient of determination of 23%, the variable can explain 23% of the variance. This value is closely followed by the correlation between the number of project members

and y6. Its coefficient of determination can also explain 23%. The last unforeseen correlation is the one between revenue and y6. As mentioned earlier, this could just make sense when regarding the turnover as an indicator of the project size. But here again, there seems to be no obvious connection between turnover and expenditure for licence.

The fifth strongest correlation appears between y6 and the number of users. As mentioned above, this connection makes perfect sense but would have been expected to be much stronger.

A weak relationship is discovered for the following three variables: number of organisational units or departments, finance module, and maturity of processes.
The connection to the variable “finance module” (x12-8), which might cause higher licence costs, is an interesting result.
As for the other variables, again, their significance is surprising since there is no obvious connection between them and the licence costs, but they seem to be useful indicators influencing the costs.

Regarding each coefficient of determination individually shows that a huge proportion of unexplained variance remains. It is not possible to prove if the variance could be explained by the other significant cost drivers or if completely different aspects are needed.

An overview of the results of this cost field is presented in the table below:

Table 6.30: Findings regarding licence costs during ERP evaluation phase (y6)

	Correlation (r)	**r^2**	**Correlation Strength**
Positive correlations			
x_{17} No. of external consultants	0.519	27%	moderate correlation
x_1 No. of locations	0.479	23%	moderate correlation
x_{16} No. of internal project members	0.477	23%	moderate correlation
x_4 Revenue	0.464	22%	moderate correlation
x_{10} No. of total users	0.440	19%	moderate correlation
x2 No. of org. units or depts.	0.314	10%	low correlation
x12-8 Finance module	0.313	10%	low correlation
Negative Correlations			
x31 Maturity of processes	-0.347	12%	moderate correlation

HARDWARE COSTS (Y7)

The correlation analysis for the cost field "hardware costs" (y7) discovered 6 cost drivers at the significance level 0.01 which correlate significantly with y7. This is the first cost field which shows only positive significant correlations. Among them, five cost drivers show moderate correlations, and only one features a low correlation.

The ones with the moderate correlation are "number of organisational units or departments", "ERP system – SAP", "number of internal project managers", "revenue", "number of total users" and "number of external experts".

An overview of the correlating cost drivers arranged in a descending order sorted by their correlation coefficient from +1 to -1 is presented in the table below.

Table 6.31: Correlation between y7 and cost drivers during implementation phase

Independent Variables: Cost Drivers	N	Pearson Corr.	r^2	Signif. 2-tailed (p-Value)	Significance level	Interpretation
x_2 No. of org. units or depts.	72	0.571	33%	0	**	moderate correlation
$x_{5\text{-}1}$ ERP system - SAP	72	0.441	19%	0	**	moderate correlation
x_{16} No. of internal project members	70	0.431	19%	0	**	moderate correlation
x_4 Revenue	69	0.426	18%	0	**	moderate correlation
x_{10} No. of total users	72	0.415	17%	0	**	moderate correlation
x_{17} No. of external consultants	70	0.385	15%	0.001	**	low correlation
x_{11} No. of user groups	71	0.236	6%	0.048	*	low correlation
$x_{12\text{-}8}$ Finance module	72	0.155	2%	0.192		slight correlation
x_1 No. of locations	72	0.149	2%	0.212		slight correlation
$x_{12\text{-}5}$ DMS module	72	0.146	2%	0.221		slight correlation
x_{21} Team maturity	68	0.146	2%	0.233		slight correlation
x_{26} Consulting quality	68	0.142	2%	0.248		slight correlation
x_8 No. of reports	67	0.139	2%	0.263		slight correlation
x_{25} Consulting experience	67	0.126	2%	0.31		slight correlation
$x_{5\text{-}3}$ ERP system - proALPHA	72	0.125	2%	0.296		slight correlation
x_{32} Complexity of bus. processes	72	0.122	1%	0.253		slight correlation
$x_{12\text{-}12}$ SCM module	72	0.121	1%	0.31		slight correlation
$x_{12\text{-}2}$ MM module	72	0.082	1%	0.494		slight correlation
x_{20} Team quality	71	0.082	1%	0.498		slight correlation
x_{13} Complexity of data	72	0.080	1%	0.503		slight correlation
x_6 No. of interfaces	69	0.076	1%	0.535		slight correlation
$x_{12\text{-}9}$ Accounting module	72	0.068	0%	0.568		no correlation
$x_{12\text{-}11}$ Production module	72	0.065	0%	0.585		no correlation

$x_{12\text{-}3}$ Calculation module	72	0.063	0%	0.6		no correlation
$x_{12\text{-}10}$ HRM module	72	0.042	0%	0.728		no correlation
$x_{5\text{-}4}$ ERP system - SAGE	72	0.036	0%	0.764		no correlation
x_{22} Team composition	70	0.031	0%	0.797		no correlation
x_{19} Fit of system/organisation	71	0.022	0%	0.858		no correlation
x_{12} No.of modules	72	0.018	0%	0.881		no correlation
$x_{12\text{-}6}$ Sales module	72	0.010	0%	0.934		no correlation
x_{14} Complexity of interfaces	70	0.004	0%	0.974		no correlation
x_{9} No. of EDIs	70	0.000	0%	0.998		no correlation
x_{30} Management involvement	70	-0.009	0%	0.943		no correlation
$x_{12\text{-}4}$ PM module	72	-0.010	0%	0.932		no correlation
x_{28} User quality	68	-0.018	0%	0.887		no correlation
x_{34} Willingness to change	70	-0.030	0%	0.808		no correlation
$x_{12\text{-}1}$ Purchasing module	72	-0.038	0%	0.752		no correlation
x_{29} Employee involvement	70	-0.070	0%	0.565		no correlation
x_{15} Complexity of reports	66	-0.077	1%	0.538		slight correlation
$x_{5\text{-}5}$ ERP system - ABAS	72	-0.081	1%	0.498		slight correlation
x_{7} No. of modifications	68	-0.094	1%	0.443		slight correlation
$x_{12\text{-}13}$ Detailed planning module	72	-0.096	1%	0.423		slight correlation
x_{23} Availability of management	69	-0.096	1%	0.433		slight correlation
x_{36} Satisfaction with ERP system	71	-0.102	1%	0.395		slight correlation
x_{18} Ratio internal/external	69	-0.110	1%	0.358		slight correlation
$x_{5\text{-}2}$ ERP system - Microsoft	72	-0.116	1%	0.334		slight correlation
$x_{12\text{-}15}$ MDA module	72	-0.139	2%	0.245		slight correlation
$x_{12\text{-}7}$ CRM module	72	-0.158	2%	0.186		slight correlation
x_{24} Availability of business users	68	-0.163	3%	0.185		slight correlation
x_{31} Maturity of processes	69	-0.178	3%	0.144		slight correlation
x_{27} Critical attitude of users	70	-0.179	3%	0.139		slight correlation
$x_{12\text{-}14}$ PDA module	72	-0.212	4%	0.073		low correlation
x_{35} Commitment management	71	-0.227	5%	0.056		low correlation
$x_{5\text{-}6}$ ERP system - other	72	-0.276	8%	0.019	*	low correlation
x_{33} Stability of organisation	70	-0.284	8%	0.017	*	low correlation

As with the previous results, some results are obvious and make perfect sense, while some of the other identified relationships are hard to interpret.

The variables which are expected to have a big influence on the hardware costs are “number of locations”, “number of users” and “complexity of data”. However, the correlation between “number of users” and y7 only has a coefficient of 0.415 and a coefficient of determination of about 17%. This was bound to be higher. The other two cost drivers do not feature a significant connection at all. This seems paradox.

The strongest connection is identified between "number of organisational units or departments" and y7, which is 0.571. The coefficient of determination can predict 33% of the variance and provides a moderate guide for assessing the dependent variable field "hardware costs". Having increased hardware costs when raising the number of units or departments appears to have no obvious connection; its explanation would be pure speculation. The same seems to apply to the variables "number of internal project managers", "revenue" and "number of external consultants".

The connections to "number of users", which is mentioned above, and the relationship to "ERP system-SAP" are more reasonable. The correlation between SAP-vendor and y7 is 0.441 and thus the second strongest connection. Its coefficient of determination predicts 19% of the variance. This moderate correlation indicates that companies choosing SAP are likely to spend more on hardware than on other ERP providers. This seems to be an interesting result.

At this point, as in the previous section and chapters, one can observe a remaining quantity of unexplained variance when looking at each correlation individually. It cannot be assessed at this point in the study if the remaining variance could be explained by the other significant cost drivers.

To summarise this section, the 7 identified cost drivers influencing hardware costs are listed below:

Table 6.32: Findings regarding hardware costs during ERP evaluation phase (y7)

	Correlation (r)	r^2	Corrrelation Strength
Positive correlations			
x_2 No. of org. units or depts.	0.571	33%	moderate correlation
$x_{5\text{-}1}$ ERP system - SAP	0.441	19%	moderate correlation
x_{16} No. of internal project members	0.431	19%	moderate correlation
x_4 Revenue	0.426	18%	moderate correlation
x_{10} No. of total users	0.415	17%	moderate correlation
x_{17} No. of external consultants	0.385	15%	low correlation

6.3.3 RQ2-3: ANALYSIS & FINDINGS

This study has already presented the correlation analysis for the cost fields during the evaluation and implementation phase.

Now, each cost field of the evaluation phase (internal personnel costs, external personnel costs, ERP software costs, licence costs, and hardware costs) should be analysed for its correlation with the potential cost drivers. As in the previous chapters, the cost fields will be presented successively.

INTERNAL PERSONNEL COSTS (Y8) | MAINTENANCE

At the significance level 0.01, the analysis identified a correlation between 13 cost drivers and the internal personnel costs during the maintenance phase (y8).

A positive connection is discovered for the following ones: “number of total users”, “number of units or departments”, “number of external consultants”, “number of locations”, “revenue”, “ERP System –SAP”, “number of electronical data interfaces (EDIs)”, “number of project members”, “number of interfaces”, “supply chain management module” and “number of reports”.

A negative relationship could be identified for the variables “ERP system – other” and “stability of organisation”.

An overview of the correlating cost drivers arranged in a descending order sorted by their correlation coefficient from +1 to -1 is presented in the table below.

Table 6.33: Correlation between y8 and cost drivers during maintenance phase

Independent Variables: Cost Drivers	N	Pearson Corr.	r²	Signif. 2-tailed (p-Value)	Significance level	Interpretation
x_{10} No. of total users	72	0.798	64%	0	**	high correlation
x_2 No. of org. units or depts.	72	0.627	39%	0	**	moderate correlation
x_{17} No. of external consultants	70	0.573	33%	0	**	moderate correlation
x_1 No. of locations	72	0.533	28%	0	**	moderate correlation
x_4 Revenue	69	0.472	22%	0	**	moderate correlation
x_{5-1} ERP system - SAP	72	0.468	22%	0	**	moderate correlation
x_9 No. of EDIs	70	0.332	11%	0.005	**	low correlation
x_{16} No. of internal project members	70	0.325	11%	0.006	**	low correlation

x_6 No. of interfaces	69	0.324	10%	0.007	**	low correlation
$x_{12\text{-}12}$ SCM module	72	0.324	10%	0.005	**	low correlation
x_8 No. of reports	67	0.323	10%	0.008	**	low correlation
x_{12} No. of modules	72	0.289	8%	0.014	*	low correlation
$x_{12\text{-}5}$ DMS module	72	0.275	8%	0.02	*	low correlation
$x_{12\text{-}8}$ Finance module	72	0.266	7%	0.024	*	low correlation
x_{13} Complexity of data	72	0.230	5%	0.051		low correlation
$x_{12\text{-}3}$ Calculation module	72	0.218	5%	0.066		low correlation
$x_{12\text{-}9}$ Accounting module	72	0.218	5%	0.065		low correlation
$x_{12\text{-}10}$ HRM module	72	0.168	3%	0.157		slight correlation
$x_{12\text{-}11}$ Production module	72	0.166	3%	0.164		slight correlation
$x_{12\text{-}4}$ PM module	72	0.152	2%	0.204		slight correlation
x_{18} Ratio internal/external	69	0.131	2%	0.28		slight correlation
$x_{12\text{-}1}$ Purchasing module	72	0.107	1%	0.373		slight correlation
x_{14} Complexity of interfaces	70	0.094	1%	0.44		slight correlation
x_{21} Team maturity	68	0.091	1%	0.46		slight correlation
$x_{12\text{-}6}$ Sales module	72	0.087	1%	0.466		slight correlation
$x_{12\text{-}7}$ CRM module	72	0.071	1%	0.555		slight correlation
$x_{5\text{-}2}$ ERP system - Microsoft	72	0.067	0%	0.577		no correlation
x_{28} User quality	68	0.063	0%	0.611		no correlation
x_{11} No. of user groups	71	0.060	0%	0.617		no correlation
x_{23} Availability of management	69	0.048	0%	0.694		no correlation
x_7 No. of modifications	78	0.042	0%	0.731		no correlation
x_{15} Complexity of reports	66	0.041	0%	0.744		no correlation
x_{20} Team quality	71	0.039	0%	0.744		no correlation
$x_{12\text{-}2}$ MM module	72	0.038	0%	0.754		no correlation
x_{36} Satisfaction with ERP system	71	0.036	0%	0.767		no correlation
x_{32} Complexity of bus. processes	72	-0.006	0%	0.888		no correlation
x_{29} Employee involvement	70	-0.028	0%	0.817		no correlation
x_{34} Willingness to change	70	-0.029	0%	0.809		no correlation
$x_{5\text{-}3}$ ERP system - proALPHA	72	-0.033	0%	0.78		no correlation
$x_{5\text{-}5}$ ERP system - ABAS	72	-0.051	0%	0.67		no correlation
x_{26} Consulting quality	68	-0.052	0%	0.671		no correlation
x_{25} Consulting experience	67	-0.054	0%	0.666		no correlation
$x_{5\text{-}4}$ ERP system - SAGE	72	-0.060	0%	0.616		no correlation
x_{19} Fit of system/organisation	71	-0.071	1%	0.557		slight correlation
x_{22} Team composition	70	-0.110	1%	0.363		slight correlation
x_{27} Critical attitude of users	70	-0.117	1%	0.336		slight correlation
$x_{12\text{-}15}$ MDA module	72	-0.132	2%	0.271		slight correlation
$x_{12\text{-}13}$ Detailed planning module	72	-0.182	3%	0.126		slight correlation
$x_{12\text{-}14}$ PDA module	72	-0.184	3%	0.121		slight correlation
x_{35} Commitment management	71	-0.185	3%	0.123		slight correlation
x_{24} Availability of business users	68	-0.229	5%	0.061		low correlation

x_{31} Maturity of processes	69	-0.238	6%	0.049	*	low correlation
x_{30} Management involvement	70	-0.275	8%	0.021	*	low correlation
$x_{5\text{-}6}$ ERP system - other	72	-0.315	10%	0.007	**	low correlation
x_{33} Stability of organisation	70	-0.361	13%	0.002	**	low correlation

The table shows 1 high correlation, 5 moderate ones and 7 low connections. Among them, some correlations are obvious and some are astonishing. Other variables, like "number of modules", "complexity of reports" or "complexity of business processes" would have been expected to provide at least a moderate guide to assess the cost field "internal personnel costs" during the maintenance phase, but some of them do not show a connection at all.

The cost driver "total number of users" (x10) shows the strongest relationship. The correlation between this variable and the internal personnel costs during the maintenance phase is 0.798. With a coefficient of determination of 64%, this variable delivers a very good guide for assessing y8. This correlation was expected since an increased number of users increases the number of internal staff maintaining the ERP system.

Another anticipated relationship is the highest moderate correlation which is discovered between y8 and the "number of organisational units or departments". The correlation between them is +0.627 and its coefficient of determination shows that the quantity of units or departments can explain 39% of the variance.

The positive relationship seems to be very logical. The modules of each department need to be maintained at least by internal personnel. This appears to occupy the personnel and causes costs.

This argument is also relevant for the positive relationship between y8 and the "number of locations". With a value of 0.533, this independent variable is the third strongest among the moderate correlations and able to predict 28% of the variance of the required costs for internal personnel during the maintenance phase. The internal personnel are at least co-responsible for keeping the system in good condition in many companies. Thus it seems plausible that more locations require more maintenance and consequently generate higher costs for internal personnel.

Another group of cost drivers is indicative of the issue of complexity. These are the following three variables: "number of EDIs", "number of interfaces" and "number of reports". With coefficients of determination between 10% and 11%, all of them have low correlations but imply an important fact. If there are more interfaces and reports, more needs to be kept in good condition. It appears logical to have an increase of costs for internal personnel when having a high quantity of EDIs, interfaces and reports.

In this context, the positive correlation between "number of internal project members", which features a value of 0.325 and a coefficient of determination of 11%, appears very obvious. The more members are involved in the maintenance process, the higher the costs for it. That makes perfect sense and does not need any further explanations.

The relationship between "ERP system – SAP" and y8 is interesting. With a correlation of 0.465 and a coefficient of determination of 22%, it is the lowest moderate relationship.
This indicates that organisations choosing SAP have higher costs for internal personnel during the maintenance phase. In this context, the negative correlation of -0.315 and a coefficient of determination of 10% is remarkable. This implies that other vendors (the ones beside the famous ones, like Microsoft, proALPHA, SAGE, and ABAS which receive their own sub-variable) cause lower costs for internal personnel at this stage of the lifecycle.

There are two unforeseen correlations: the positive correlation between "number of external consultants" and the negative one to "stability of organisation".
An astonishing result is the relatively strong positive connection to "external consultants" which was also identified for y3. As mentioned at that point, at first glance, the connection seems to be rather contradictory and inconsistent. Possible explanations are given at y3.
At least the miracle of the variable "stability of organisation" appears again at this stage. Possible explanations are given in RQ1-1 and can be transferred to this section.

Finalising this chapter, an overview of the findings is presented in the table 6.34.

Table 6.34: Findings internal personnel costs during ERP maintenance phase (y8)

	Correlation (r)	r²	Corrrelation Strength
Positive Correlations			
x_{10} No. of total users	0.798	64%	high correlation
x_2 No. of org. units or depts.	0.627	39%	moderate correlation
x_{17} No. of external consultants	0.573	33%	moderate correlation
x_1 No. of locations	0.533	28%	moderate correlation
x_4 Revenue	0.472	22%	moderate correlation
x_{5-1} ERP system - SAP	0.468	22%	moderate correlation
x_9 No. of EDIs	0.332	11%	low correlation
x_{16} No. of internal project members	0.325	11%	low correlation
x_6 No. of interfaces	0.324	10%	low correlation
x_{12-12} SCM module	0.324	10%	low correlation
x_8 No. of reports	0.323	10%	low correlation
Negative Correlations			
x5-6 ERP system - other	-0.315	10%	low correlation
x33 Stability of organisation	-0.361	13%	low correlation

EXTERNAL PERSONNEL COSTS (Y9) | MAINTENANCE

Eleven variables were identified to have a relation to the external personnel costs during the maintenance phase at the significance level 0.01.

The correlation analysis for this cost field discovered 9 positive connections, which are: “number of total users”, “number of external consultants”, “number of locations”, “number of internal project managers”, “revenue”, number of units or departments”, “ERP system – SAP”, “number of modules”, and “DMS module”.

Negative correlations were found between the cost driver “stability of organisation” as well as “ERP system – other” and y9.

Table 6.35 gives an overview of the correlating cost drivers arranged in a descending order and sorted by their correlation coefficient from +1 to -1.

Table 6.35: Correlation between y9 and cost drivers during maintenance phase

Independent Variables: Cost Drivers	N	Pearson Corr.	r²	Signif. 2-tailed (p-Value)	Significance level	Interpretation
x_{10} No. of total users	72	0.778	61%	0	**	high correlation
x_{17} No. of external consultants	70	0.736	54%	0	**	high correlation

x_1 No. of locations	72	0.560	31%	0	**	moderate correlation
x_{16} No. of internal project members	70	0.449	20%	0	**	moderate correlation
x_4 Revenue	69	0.397	16%	0.001	**	low correlation
x_2 No. of org. units or depts.	72	0.392	15%	0.001	**	low correlation
$x_{5\text{-}1}$ ERP system - SAP	72	0.391	15%	0.001	**	low correlation
x_{12} No. of modules	72	0.371	14%	0.001	**	low correlation
$x_{12\text{-}5}$ DMS module	72	0.323	10%	0.006	**	low correlation
x_6 No. of interfaces	69	0.299	9%	0.013	*	low correlation
$x_{12\text{-}8}$ Finance module	72	0.299	9%	0.011	*	low correlation
x_9 No. of EDIs	70	0.277	8%	0.02	*	low correlation
$x_{12\text{-}12}$ SCM module	72	0.272	7%	0.021	*	low correlation
$x_{12\text{-}9}$ Accounting module	72	0.260	7%	0.027	*	low correlation
$x_{12\text{-}3}$ Calculation module	72	0.251	6%	0.034	*	low correlation
x_8 No. of reports	67	0.246	6%	0.045	*	low correlation
$x_{12\text{-}4}$ PM module	72	0.239	6%	0.044	*	low correlation
$x_{12\text{-}10}$ HRM module	72	0.231	5%	0.05		low correlation
x_{13} Complexity of data	72	0.221	5%	0.063		low correlation
$x_{5\text{-}2}$ ERP system - Microsoft	72	0.219	5%	0.064		low correlation
x_{14} Complexity of interfaces	70	0.176	3%	0.144		slight correlation
$x_{12\text{-}11}$ Production module	72	0.143	2%	0.229		slight correlation
$x_{12\text{-}1}$ Purchasing module	72	0.077	1%	0.518		slight correlation
x_7 No. of modifications	68	0.064	0%	0.604		no correlation
x_{29} Employee involvement	70	0.062	0%	0.611		no correlation
$x_{12\text{-}2}$ MM module	72	0.058	0%	0.63		no correlation
$x_{12\text{-}6}$ Sales module	72	0.056	0%	0.638		no correlation
x_{15} Complexity of reports	66	0.054	0%	0.667		no correlation
$x_{12\text{-}15}$ MDA module	72	0.024	0%	0.842		no correlation
$x_{12\text{-}7}$ CRM module	72	0.005	0%	0.967		no correlation
x_{18} Ratio internal/external	69	0.004	0%	0.965		no correlation
x_{21} Team maturity	68	0.002	0%	0.989		no correlation
$x_{5\text{-}5}$ ERP system - ABAS	72	-0.002	0%	0.983		no correlation
x_{23} Availability of management	69	-0.013	0%	0.918		no correlation
$x_{5\text{-}3}$ ERP system - proALPHA	72	-0.016	0%	0.892		no correlation
x_{20} Team quality	71	-0.020	0%	0.87		no correlation
x_{27} Critical attitude of users	70	-0.050	0%	0.683		no correlation
x_{36} Satisfaction with ERP system	71	-0.051	0%	0.675		no correlation
x_{28} User quality	68	-0.056	0%	0.648		no correlation
x_{11} No. of user groups	71	-0.058	0%	0.632		no correlation
$x_{5\text{-}4}$ ERP system - SAGE	72	-0.074	1%	0.535		slight correlation
x_{26} Consulting quality	68	-0.099	1%	0.424		slight correlation
$x_{12\text{-}13}$ Detailed planning module	72	-0.120	1%	0.315		slight correlation
$x_{12\text{-}14}$ PDA module	72	-0.127	2%	0.289		slight correlation

x_{30} Management involvement	70	-0.128	2%	0.289		slight correlation
x_{32} Complexity of bus. processes	72	-0.139	2%	0.221		slight correlation
x_{25} Consulting experience	67	-0.146	2%	0.238		slight correlation
x_{19} Fit of system/organisation	71	-0.152	2%	0.206		slight correlation
x_{24} Availability of business users	68	-0.160	3%	0.193		slight correlation
x_{34} Willingness to change	70	-0.167	3%	0.168		slight correlation
x_{22} Team composition	70	-0.189	4%	0.116		slight correlation
x_{31} Maturity of processes	69	-0.282	8%	0.019	*	low correlation
x_{35} Commitment management	71	-0.301	9%	0.011	*	low correlation
x_{33} Stability of organisation	70	-0.333	11%	0.005	**	low correlation
x_{5-6} ERP system - other	72	-0.404	16%	0	**	moderate correlation

The table shows 2 high correlations, 3 moderate ones and 6 low ones.
The three strongest correlations seem have an understandable relationship: The strongest connection is discovered for “number of total user”. With a correlation coefficient of 0.778, its coefficient of determination explains 61% of the variance. This value is followed by the second highest correlation, which is “number of external consultants” and “external personnel costs”. The correlation coefficient is 0.736, and its coefficient of determination explains 54% of the variance. The third strongest connection (the highest moderate one) is discovered for “number of locations”. With a correlation coefficient of 0.560, its coefficient of determination explains 31% of the variance.
These three connections appear to provide a good guide for assessing the cost field “external personnel” during the maintenance phase.

The variables “number of units or departments” and “number of modules” show only a low correlation, but their connections seem obvious.
Having more modules means that more maintenance is required to keep them in good condition. Besides internal personnel, external staff appears to be required for this job.

The next two identified relationships do not have such an obvious connection. These are the relationships between the variable “number of internal project members” and “revenue”. The correlation coefficient between the “number of internal project members” and y9 is 0.449. With a coefficient of determination of 20%, this is the cost driver with the lowest moderate connection.
With a correlation coefficient of 0.397, the variable “revenue” is the strongest among the low relationships. Its coefficient of determination explains 16% of the variance.

Both variables exist within similar contexts and possible interpretation opportunities were already given at these stages.
At this cost field again, the analysis discovered a correlation between the chosen vendor and the amount of external personnel costs. This connection was identified before at the cost fields y3, y4 and y8. Like with those independent variables, the cost driver “ERP system – other” shows a negative relationship. This finding implies that SAP causes higher expenses for external personnel than other vendors do. Having a smaller vendor (one besides the famous ones, like Microsoft, proALPHA, SAGE, and ABAS which receive their own sub-variable) seems to decrease the costs for external personnel.

The relationship between the DMS module and the expense for external personnel is interesting. This module appears to cause service intensity by external personnel than other modules. Nevertheless, with a correlation index of .323 (10%), this influence is relatively low.
In addition to the variable “ERP system – other”, the variable “stability of organisation” also shows a low negative correlation.

So far, this chapter has identified 11 cost drivers which have a relevant connection to the external personnel costs during the maintenance phase.
The table 6.36 summarises them.

Table 6.36: Findings external personnel costs during ERP maintenance phase (y9)

	Correlation (r)	r²	Correlation Strength
Positive Correlation			
x_{10} No. of total users	0.778	61%	high correlation
x_{17} No. of external consultants	0.736	54%	high correlation
x_1 No. of locations	0.560	31%	moderate correlation
x_{16} No. of internal project members	0.449	20%	moderate correlation
x_4 Revenue	0.397	16%	low correlation
x_2 No. of org. units or depts.	0.392	15%	low correlation
$x_{5\text{-}1}$ ERP system - SAP	0.391	15%	low correlation
x_{12} No. of modules	0.371	14%	low correlation
$x_{12\text{-}5}$ DMS module	0.323	10%	low correlation
Negative Correlation			
x33 Stability of organisation	-0.333	11%	low correlation
x5-6 ERP system - other	-0.404	16%	moderate correlation

ERP SOFTWARE COSTS (Y10)

After identifying more than 10 relevant cost drivers in each of the previous costs fields during the maintenance phase, the correlation analysis for y10 identified just 3 relationships between the expenses for software costs and the cost drivers at the significance level 0.001. These are “number of external consultants”, “number of project managers” and “ERP system – SAP”.

The following table gives an overview of the correlating cost drivers arranged in a descending order sorted by their correlation coefficients from +1 to -1:

Table 6.37: Correlation between y10 and cost drivers during maintenance phase

Independent Variables: Cost Drivers	N	Pearson Corr.	r²	Signif.. 2-tailed (p-Value)	Significance level	Interpretation
x_{17} No. of external consultants	70	0.442	20%	0	**	moderate correlation
x_{16} No. of internal project members	70	0.368	14%	0.002	**	low correlation
$x_{5\text{-}1}$ ERP system - SAP	72	0.314	10%	0.007	**	low correlation
x_2 No. of org. units or depts.	72	0.270	7%	0.022	*	low correlation
x_4 Revenue	69	0.239	6%	0.048	*	low correlation
x_{14} Complexity of interfaces	70	0.232	5%	0.053		low correlation
$x_{12\text{-}5}$ DMS module	72	0.210	4%	0.077		low correlation
x_{10} No. of total users	72	0.187	3%	0.117		slight correlation

$x_{5\text{-}3}$ ERP system - proALPHA	72	0.164	3%	0.17		slight correlation
$x_{12\text{-}10}$ HRM module	72	0.154	2%	0.197		slight correlation
x_{13} Complexity of data	72	0.151	2%	0.206		slight correlation
x_{30} Management involvement	70	0.148	2%	0.221		slight correlation
x_{15} Complexity of reports	66	0.146	2%	0.243		slight correlation
x_{36} Satisfaction with ERP system	71	0.132	2%	0.273		slight correlation
$x_{12\text{-}8}$ Finance module	72	0.108	1%	0.366		slight correlation
$x_{12\text{-}3}$ Calculation module	72	0.107	1%	0.373		slight correlation
x_1 No. of locations	72	0.089	1%	0.459		slight correlation
x_{32} Complexity of bus. processes	72	0.089	1%	0.384		slight correlation
$x_{12\text{-}4}$ PM module	72	0.073	1%	0.542		slight correlation
x_{20} Team quality	71	0.073	1%	0.545		slight correlation
x_{21} Team maturity	68	0.069	0%	0.577		no correlation
x_{12} No. of modules	72	0.041	0%	0.735		no correlation
$x_{12\text{-}9}$ Accounting module	72	0.041	0%	0.73		no correlation
x_{29} Employee involvement	70	0.040	0%	0.744		no correlation
x_{27} Critical attitude of users	70	0.038	0%	0.754		no correlation
$x_{12\text{-}2}$ MM module	72	0.037	0%	0.756		no correlation
x_{11} No. of user groups	71	0.031	0%	0.796		no correlation
$x_{5\text{-}5}$ ERP system - ABAS	72	0.023	0%	0.85		no correlation
x_{19} Fit of system/organisation	71	0.023	0%	0.851		no correlation
x_{18} Ratio internal/external	69	0.017	0%	0.881		no correlation
$x_{12\text{-}15}$ MDA module	72	0.015	0%	0.902		no correlation
x_{26} Consulting quality	68	0.002	0%	0.988		no correlation
x_6 No. of interfaces	69	-0.018	0%	0.881		no correlation
$x_{12\text{-}1}$ Purchasing module	72	-0.021	0%	0.862		no correlation
x_{23} Availability of management	69	-0.038	0%	0.757		no correlation
x_{31} Maturity of processes	69	-0.055	0%	0.655		no correlation
x_9 No. of EDIs	70	-0.063	0%	0.605		no correlation
x_{24} Availability of business users	68	-0.067	0%	0.59		no correlation
x_7 No. of modifications	68	-0.071	1%	0.563		slight correlation
x_8 No. of reports	67	-0.071	1%	0.568		slight correlation
$x_{12\text{-}11}$ Production module	72	-0.071	1%	0.554		slight correlation
x_{22} Team composition	70	-0.073	1%	0.549		slight correlation
$x_{12\text{-}14}$ PDA module	72	-0.083	1%	0.489		slight correlation
x_{25} Consulting experience	67	-0.087	1%	0.484		slight correlation
x_{34} Willingness to change	70	-0.088	1%	0.47		slight correlation
$x_{12\text{-}13}$ Detailed planning module	72	-0.089	1%	0.459		slight correlation
$x_{12\text{-}7}$ CRM module	72	-0.099	1%	0.41		slight correlation
$x_{5\text{-}4}$ ERP system - SAGE	72	-0.104	1%	0.387		slight correlation
x_{28} User quality	68	-0.104	1%	0.397		slight correlation
$x_{12\text{-}12}$ SCM module	72	-0.118	1%	0.322		slight correlation
x_{33} Stability of organisation	70	-0.125	2%	0.304		slight correlation
$x_{5\text{-}2}$ ERP system - Microsoft	72	-0.130	2%	0.275		slight correlation
$x_{12\text{-}6}$ Sales module	72	-0.150	2%	0.208		slight correlation
$x_{5\text{-}6}$ ERP system - other	72	-0.176	3%	0.139		slight correlation
x_{35} Commitment management	71	-0.267	7%	0.024	*	low correlation

The table shows a huge remainder of unexplained variance for the expenses of ERP software costs during the maintenance phase. Maybe other factors, which are not considered, would be more suitable to explain the quantity of costs for software during the maintenance phase.

One would have expected cost drivers like “number of users”, “number of modules”, “willingness to change” and “complexity of reports”, but none of them show a significant impact on the cost field y10.

At any rate, 20% of the variance can be explained by the “number of external consultants” which provides a moderate guide for assessing the cost field “software costs”.

This value is followed by the two remaining variables. The positive correlation between “number of internal project members” and y10 is 0.364. Its coefficient of correlation is able to predict 14% of the variance.

These two findings are very implausible. Why should the expenditure for software increase with a higher number of internal project members and external consultants. This relationship is a bit difficult to understand.

The third finding once more identified the vendor SAP as the provider effecting higher costs than other vendors. In fact, there is only a low correlation of 0.314 between SAP and the costs for software, but companies choosing it appear to expect more costs for software during the maintenance phase.

The findings are summarised in table 6.38.

Table 6.38: Findings ERP software costs during ERP evaluation phase (y10)

	Correlation (r)	r^2	Correlation Strength
Positive Correlation			
x_{17} No. of external consultants	0.442	20%	moderate correlation
x_{16} No. of internal project members	0.368	14%	low correlation
$x_{5\text{-}1}$ ERP system - SAP	0.314	10%	low correlation

LICENCE COSTS (Y11) | MAINTENANCE

Eleven variables were identified as having a significant correlation with the licence expense during the maintenance phase at the 0.01 significance level.

These are: “number of total users”, “number of external consultants”, “revenue”, “number of locations”, “number of departments or units”, number of EDIs”, “DMS module”, “number of modules”, “number of user groups” and “accounting module”.

A positive direction was discovered for all of them. The table below gives a detailed overview of the results arranged in a descending order and sorted by their correlation co-efficient from +1 to -1:

Table 6.39: Correlation between y11 and cost drivers during maintenance phase

Independent Variables: Cost Drivers	N	Pearson Corr.	r^2	Signif. 2-tailed (p-Value)	Signifi-cance level	Interpretation
x_{10} No. of total users	72	0.655	43%	0	**	moderate correlation
x_{17} No. of external consultants	70	0.548	30%	0	**	moderate correlation
x_{4} Revenue	69	0.522	27%	0	**	moderate correlation
x_{1} No. of locations	72	0.446	20%	0	**	moderate correlation
x_{2} No. of org. units or depts.	72	0.365	13%	0.002	**	low correlation
x_{16} No. of internal project members	70	0.359	13%	0.002	**	low correlation
x_{9} No. of EDIs	70	0.353	12%	0.003	**	low correlation
x_{12-5} DMS module	72	0.336	11%	0.004	**	low correlation
x_{12} No. of modules	72	0.324	10%	0.005	**	low correlation
x_{11} No. of user groups	71	0.317	10%	0.007	**	low correlation
x_{12-9} Accounting module	72	0.306	9%	0.009	**	low correlation
x_{12-10} HRM module	72	0.291	8%	0.013	*	low correlation
x_{12-8} Finance module	72	0.266	7%	0.024	*	low correlation
x_{13} Complexity of data	72	0.261	7%	0.027	*	low correlation
x_{5-1} ERP system - SAP	72	0.232	5%	0.049	*	low correlation
x_{6} No. of interfaces	69	0.206	4%	0.089		low correlation
x_{12-3} Calculation module	72	0.203	4%	0.087		low correlation
x_{8} No. of reports	67	0.191	4%	0.121		slight correlation
x_{12-4} PM module	72	0.181	3%	0.127		slight correlation
x_{21} Team maturity	68	0.162	3%	0.186		slight correlation
x_{14} Complexity of interfaces	70	0.149	2%	0.22		slight correlation
x_{5-2} ERP system - Microsoft	72	0.131	2%	0.273		slight correlation
x_{12-12} SCM module	72	0.131	2%	0.273		slight correlation
x_{29} Employee involvement	70	0.130	2%	0.282		slight correlation
x_{20} Team quality	71	0.112	1%	0.353		slight correlation
x_{12-2} MM module	72	0.110	1%	0.359		slight correlation
x_{18} Ratio internal/external	69	0.108	1%	0.386		slight correlation
x_{12-7} CRM module	72	0.104	1%	0.387		slight correlation
x_{5-3} ERP system - proALPHA	72	0.098	1%	0.412		slight correlation
x_{32} Complexity of bus. processes	72	0.075	1%	0.361		slight correlation
x_{12-11} Production module	72	0.067	0%	0.577		no correlation
x_{22} Team composition	70	0.060	0%	0.621		no correlation
x_{12-14} PDA module	72	0.043	0%	0.723		no correlation
x_{24} Availability of business users	68	0.037	0%	0.767		no correlation
x_{19} Fit of system/organisation	71	0.036	0%	0.763		no correlation
x_{15} Complexity of reports	66	0.031	0%	0.805		no correlation

x_{36} Satisfaction with ERP system	71	0.001	0%	0.996		no correlation
x_{23} Availability of management	69	-0.013	0%	0.915		no correlation
x_{25} Consulting experience	67	-0.016	0%	0.901		no correlation
x_{26} Consulting quality	68	-0.027	0%	0.826		no correlation
x_{7} No. of modifications	68	-0.035	0%	0.778		no correlation
$x_{12\text{-}1}$ Purchasing module	72	-0.035	0%	0.77		no correlation
x_{34} Willingness to change	70	-0.058	0%	0.631		no correlation
$x_{12\text{-}6}$ Sales module	72	-0.063	0%	0.598		no correlation
$x_{5\text{-}5}$ ERP system - ABAS	72	-0.082	1%	0.492		slight correlation
x_{28} User quality	68	-0.103	1%	0.405		slight correlation
x_{30} Management involvement	70	-0.136	2%	0.261		slight correlation
$x_{12\text{-}15}$ MDA module	72	-0.142	2%	0.235		slight correlation
$x_{12\text{-}13}$ Detailed planning module	72	-0.151	2%	0.206		slight correlation
$x_{5\text{-}4}$ ERP system - SAGE	72	-0.163	3%	0.171		slight correlation
x_{27} Critical attitude of users	70	-0.173	3%	0.152		slight correlation
x_{35} Commitment management	71	-0.193	4%	0.107		slight correlation
$x_{5\text{-}6}$ ERP system - other	72	-0.199	4%	0.094		slight correlation
x_{31} Maturity of processes	69	-0.248	6%	0.04	*	low correlation
x_{33} Stability of organisation	70	-0.291	8%	0.015	*	low correlation

The table shows four moderate relationships:

The strongest one is “number of total users” with a correlation coefficient of -0.655. The number of users appears to provide only a moderate guide to the licence costs as it predicts 43% of the variance. The identified relationship indicates that the higher the number of users, the higher the licence costs.

The second strongest relationship is the “number of total external consultants” with a value of + 0.548. With a coefficient of determination of 30%, the number of consultants moderately explains the licence costs during the maintenance phase. The finding means that the higher the number of consultants, the higher the required costs for licence.

With a correlation coefficient of 0.522, the variable “revenue” has the third strongest relationship to y11. The coefficient of determination expresses that the variable explains 27% of the variance.

With a coefficient of determination of 20%, the variable “number of locations” represents the fourth moderate correlation. The correlation between this cost driver and the licence costs during the maintenance phase is 0.446.

The 7 remaining cost drivers “number of departments or units”, “number of internal project managers”, “number of EDIs”, “DMS module”, “number of modules”, “number of user groups” and “accounting module” feature a significance at the 0.01 level but show only a low positive correlation.

All other cost drivers do not have any or only slight correlations.

So far, this chapter has identified 11 relevant relationships between the cost drivers and licence costs. These are summarised in table 6.40.

Table 6.40: Findings licence costs during ERP maintenance phase (y11)

	Correlation (r)	r^2	Correlation Strength
Positive Correlations			
x_{10} No. of total users	0.655	43%	moderate correlation
x_{17} No. of external consultants	0.548	30%	moderate correlation
x_4 Revenue	0.522	27%	moderate correlation
x_1 No. of locations	0.446	20%	moderate correlation
x_2 No. of org. units or depts.	0.365	13%	low correlation
x_{16} No. of internal project members	0.359	13%	low correlation
x_9 No. of EDIs	0.353	12%	low correlation
$x_{12\text{-}5}$ DMS module	0.336	11%	low correlation
x_{12} No. of modules	0.324	10%	low correlation
x_{11} No. of user groups	0.317	10%	low correlation
$x_{12\text{-}9}$ Accounting module	0.306	9%	low correlation

As quite often in this chapter, correlation analysis determined some apparent results whilst others are very unobvious.
Starting with the anticipated results, a positive correlation between number of total users and licence costs was quite expected since the licence costs usually emerge for each user. The moderate positive relationship to the number of users consequently seems to be very logical.
As with the number of users, the positive relationship to the number of modules is apparent. Since each module requires a licence, it appears to be natural that the costs increase with the number of modules. However, the influence of this variable was expected to be stronger. At the moment, it predicts only 10% of the variance and provides just a low guide for assessing the licence costs.

Another understandable result is the positive connection to "number of EDIs". Licence costs can emerge for interfaces, which makes the positive relationship between this variable and y11 very plausible.

The last comprehensible outcome is the positive connection between the two modules DMS and accounting. It is very interesting that they appear to generate higher licence costs during the maintenance phase than other modules. One explanation could be that they are more expensive than other modules and thus generate more costs for licences.

The variables "number of external personnel", "revenue", "number of locations", "number of departments or units", "number of internal project members" and "number of user groups" provided unapparent findings.

All of these variables do not have a direct or obvious connection to the amount of the licence costs, but they have one thing in common: they are all surely indicators of the size of a project, and, of course, the project size appears to influence the costs.

As mentioned earlier, this correlation analysis cannot deliver causal relationships between variables, but aims to find good indicators for licence costs.

HARDWARE COSTS (Y12) | MAINTENANCE

Last but not least, the correlation results for hardware costs during the maintenance phase will be presented.

The correlation analysis identified 4 cost drivers with moderate or low relationships at the significance level 0.01. They are: "number of external consultants", "number of internal project managers", "number of total users" and "revenue". All of these variables show a positive relationship to y12.

The following table provides an overview of the results, which are arranged in a descending order and sorted by the correlation coefficient from +1 to -1.

Table 6.41: Correlation between y12 and cost drivers during maintenance phase

Independent Variables: Cost Drivers	N	Pearson Corr.	r^2	Signif. 2-tailed (p-Value)	Signifi-cance level	Interpretation
x_{17} No. of external consultants	70	0.569	32%	0	**	moderate correlation
x_{16} No. of internal project members	70	0.428	18%	0	**	moderate correlation
x_{10} No. of total users	72	0.405	16%	0	**	moderate correlation
x_4 Revenue	69	0.312	10%	0.009	**	low correlation
x_2 No. of org. units or depts.	72	0.294	9%	0.012	*	low correlation
x_{13} Complexity of data	72	0.224	5%	0.058		low correlation
x_{5-1} ERP system - SAP	72	0.216	5%	0.068		low correlation
x_{12-3} Calculation module	72	0.216	5%	0.068		low correlation
x_{14} Complexity of interfaces	70	0.186	3%	0.122		slight correlation
x_{12-5} DMS module	72	0.173	3%	0.145		slight correlation
x_1 No. of locations	72	0.172	3%	0.15		slight correlation
x_8 No. of reports	67	0.159	3%	0.199		slight correlation
x_{15} Complexity of reports	66	0.117	1%	0.347		slight correlation
x_{26} Consulting quality	68	0.115	1%	0.352		slight correlation
x_{21} Team maturity	68	0.112	1%	0.364		slight correlation
x_{12-10} HRM module	72	0.111	1%	0.352		slight correlation
x_{20} Team quality	71	0.111	1%	0.355		slight correlation
x_{32} Complexity of bus. processes	72	0.108	1%	0.296		slight correlation
x_{25} Consulting experience	67	0.097	1%	0.436		slight correlation
x_{30} Management involvement	70	0.092	1%	0.449		slight correlation
x_{36} Satisfaction with ERP system	71	0.085	1%	0.479		slight correlation
x_{11} No. of user groups	71	0.083	1%	0.491		slight correlation
x_6 No. of interfaces	69	0.073	1%	0.551		slight correlation
x_{28} User quality	68	0.071	1%	0.563		slight correlation
x_{12-4} PM module	72	0.063	0%	0.602		no correlation
x_{5-3} ERP system - proALPHA	72	0.050	0%	0.677		no correlation
x_{12-11} Production module	72	0.049	0%	0.682		no correlation
x_{12} No. of modules	72	0.047	0%	0.695		no correlation
x_{12-2} MM module	72	0.046	0%	0.703		no correlation
x_9 No. of EDIs	70	0.044	0%	0.715		no correlation
x_{22} Team composition	70	0.038	0%	0.754		no correlation
x_{12-9} Accounting module	72	0.022	0%	0.857		no correlation
x_{19} Fit of system/organisation	71	0.014	0%	0.905		no correlation
x_{12-1} Purchasing module	72	0.007	0%	0.953		no correlation
x_{34} Willingness to change	70	-0.005	0%	0.966		no correlation
x_{27} Critical attitude of users	70	-0.014	0%	0.909		no correlation
x_{5-4} ERP system - SAGE	72	-0.022	0%	0.854		no correlation
x_{24} Availability of business users	68	-0.023	0%	0.851		no correlation
x_{12-14} PDA module	72	-0.025	0%	0.836		no correlation
x_{31} Maturity of processes	69	-0.025	0%	0.839		no correlation
x_{5-5} ERP system - ABAS	72	-0.038	0%	0.751		no correlation
x_{12-8} Finance module	72	-0.039	0%	0.747		no correlation
x_{12-12} SCM module	72	-0.049	0%	0.685		no correlation
x_{18} Ratio internal/external	69	-0.049	0%	0.682		no correlation
x_{29} Employee involvement	70	-0.051	0%	0.676		no correlation
x_{12-15} MDA module	72	-0.057	0%	0.633		no correlation
x_{12-7} CRM module	72	-0.063	0%	0.597		no correlation

x_{5-6} ERP system - other	72	-0.082	1%	0.493	slight correlation
x_7 No. of modifications	68	-0.084	1%	0.494	slight correlation
x_{12-6} Sales module	72	-0.084	1%	0.484	slight correlation
x_{5-2} ERP system - Microsoft	72	-0.092	1%	0.441	slight correlation
x_{23} Availability of management	69	-0.112	1%	0.361	slight correlation
x_{35} Commitment management	71	-0.183	3%	0.127	slight correlation
x_{12-13} Detailed planning module	72	-0.193	4%	0.105	slight correlation
x_{33} Stability of organisation	70	-0.216	5%	0.072	low correlation

Except for the variables "number of units or departments" and "ERP system –SAP", the analysis identified nearly the same cost drivers for hardware costs during the maintenance phase as for hardware costs during the implementation phase.

The strongest relationship with y12 is discovered for "number of external consultants" which provides a coefficient of 0.569. This indicator provides a moderate guide for assessing the cost field since it predicts 32%.

This cost driver is followed by the variable "number of internal project managers" with a correlation coefficient of 0.428. With 18%, its coefficient of determination is able to explain the variance modestly.

The third moderate correlation is identified for "number of users" which predicts 16% of the variance. This relationship was expected to be stronger because dependence seems obvious and makes perfect sense.

With a correlation coefficient of 0.312, the last determined relationship between y12 and the cost driver "revenue" is only a low one. This variable has explanatory power of 10% and contributes merely a low portion for assessing the hardware costs during the maintenance phase.

Some of these determined connections are rather unapparent. Only the variable "number of total users" appears to make sense at first glance. Having a higher number of users means there is a need for more computers, maybe more powerful servers to handle the data volume, and so on. The other ones seem to be more indicators of the project size which, of course, appears to have an influence on the costs. The variable "number of locations" was expected to have at least a moderate influence, but, surprisingly, it showed no significant relationship.

As mentioned before, this correlation analysis is not able to provide causal context but aims to find indicators which matter in terms of influencing the costs.

For the cost field "hardware costs during the maintenance phase", these are summarised in the following table.

Table 6.42: Findings hardware costs during ERP maintenance phase (y12)

	Correlation (r)	r^2	Correlation Strength
Positive correlations			
x_{17} No. of external consultants	0.569	32%	moderate correlation
x_{16} No. of internal project members	0.428	18%	moderate correlation
x_{10} No. of total users	0.405	16%	moderate correlation
x_4 Revenue	0.312	10%	low correlation

6.3.4 SUMMARY & CONCLUSION RQ2

RQ2 was able to answer the question which cost drivers influence ERP costs, and could create a benefit for project managers, ERP vendors and customers.

Within the literature, many cost drivers are in debate. Some were empirical tests, while others originated from theoretical frameworks or different cost estimation models and have not been empirically confirmed before. A lot of assumptions and speculation appears within the cost estimation of ERP systems.

This study aims to make a contribution to finding reliable variables or indicators which are able to access the ERP costs. Having pooled all cost drivers ever stated (identified by the systematic literature review), this chapter provides a first overall scientific view of empirically validated cost drivers. These cost drivers were tested for their significance within the sample.

As described within the conceptual framework, costs can arise for different cost fields during the evaluation, the implementation and the maintenance phase. Many studies do not state which costs they relate to and at which time they emerge. This study, by contrast, created a clear connection between cost drivers, cost fields and the lifecycle phase in which they arise. A correlation analysis for each of the cost fields was conducted.

The analysis identified significant cost drivers for each cost field at the significance level 0.01. The findings are presented by the summary table below.

Table 6.43: Findings RQ 2: Correlation of cost fields and cost drivers

Cost Fields	Correlating Cost Drivers	Strength of Correlation (r)
y1	x10 No. of total users	+0.408
Internal Personnel Costs	x1 No. of locations	+0.344
Evaluation Phase	x36 Satisfaction with ERP system	-0.309
	x31 Maturity of process	-0.358
	x35 Commitment management	-0.390
	x33 Stability of organisation	-0.422
y2	x1 No. of locations	+0.406
External Personnel Costs	x10 No. of total users	+0.399
Evaluation Phase	x21 Team maturity	-0.346
	x31 Maturity of processes	-0.361
	x36 Satisfaction with ERP system	-0.426
	x33 Stability of organisation	-0.455
	x35 Commitment management	-0.460
y4	x17 No. of external consultants	+0.719
External Personnel Costs	x10 No. of total users	+0.612
Implementation Phase	x16 No. of internal project members	+0.567
	x5-1 ERP system – SAP	+0.482
	x2 No. of organisational units or depts.	+0.443
	x12-9 Accounting module	+0.383
	x12-10 HRM module	+0.374
	x4 Revenue	+0.365
	x1 No. of locations	+0.356
	x33 Stability of organisation	-0.307
	x5-6 ERP system – other	-0.438
y5	x4 Revenue	+0.543
ERP Software Costs	x10 No. of total users	+0536
Implementation Phase	x17 No. of external consultants	+0.383
	x2 No. of organisational units or depts.	+0.346

	x1 No. of locations	+0.342
	x11 No. of user groups	+0.322
	x16 No. of internal project members	+0.322
	x33 Stability of organisation	-0.389
y6	x17 No. of external consultants	+0.519
Licence Costs	x1 No. of locations	+0.479
Implementation Phase	x16 No. of internal project members	+0.477
	x4 Revenue	+0.464
	x10 No. of total users	+0.440
	x2 No. of organisational units or depts.	+0.314
	x12-8 Finance module	+0.313
	x31 Maturity of processes	-0.347
y7	x2 No. of organisational units or depts.	+0.751
Hardware Costs	x5-1 ERP system – SAP	+0.441
Implementation Phase	x16 No. of internal project members	+0.431
	x4 Revenue	+0.426
	x10 No. of total users	+0.415
	x17 No. of external consultants	+0.385
y8	x10 No. of total users	+0.798
Internal Personnel Costs	x2 No. of organisational units or depts.	+0.627
Maintenance Phase	x17 No. of external consultants	+0,573
	x1 No. of locations	+0.533
	x4 Revenue	+0.472
	x5-1 ERP system – SAP	+0.468
	x9 No. of EDIs	+0.332
	x16 No. of internal project members	+0.325
	x6 No. of interfaces	+0.324
	x12-12 SCM module	+0.324
	x8 No. of reports	+0.323
	x5-6 ERP system – other	-0.315
	x33 Stability of organisation	-0.361
y9	x10 No. of total users	+0.778
External Personnel Costs	x17 No. of external consultants	+0.736
Maintenance Phase	x1 No. of locations	+0.560
	x16 No. of internal project members	+0.449
	x4 Revenue	+0.397
	x2 No. of organisational units or depts.	+0.392
	x5-1 ERP system – SAP	+0.391
	x12 No. of modules	+0.371
	x12-5 DMS module	+0.323
	x35 Commitment management	-0.301
	x33 Stability of organisation	-0.333
	x5-6 ERP system – other	-0.404

y10	x17 No. of external consultants	+0.442
ERP Software Costs	x16 No. of internal project members	+0.368
Maintenance Phase	x5-1 ERP system – SAP	+0.314
y11	x10 No. of total users	+0.655
Licence Costs	x17 No. of external consultants	+0.548
Maintenance Phase	x4 Revenue	+0.522
	x1 No. of locations	+0.466
	x2 No. of orgnisational units or depts.	+0.365
	x16 No. of internal project members	+0.359
	x9 No. of EDIs	+0.353
	x12-5 DMS module	+0.336
	x12 No. of modules	+0.324
	x11 No. of user groups	+0.317
	x12-9 Accounting module	+0.306
y12	x17 No. of external consultants	+0.569
Hardware Costs	x16 No. of internal project members	+0.428
Maintenance Phase	x10 No. of total users	+0.405
	x4 Revenue	+0.312

As shown in table 6.44, a total of 23 significant cost drivers could be identified; they are grouped into organisational, technical and situational ones in accordance with the categorisation of cost drivers conducted within the conceptual framework.

Table 6.44: Frequency of correlative cost drivers

Cost Driver Dimension	Cost Driver	Frequency
Organisational	x1 No. of locations	9
	x2 No. of organisational units or departments	8
	x4 Revenue	8
Technical	x5-1 ERP system - SAP	7
	x5-6 ERP system - other	4
	x6 No. of interfaces	1
	x8 No. of reports	1
	x9 No. of EDIs	2
	x10 No. of total users	11
	x11 No. of user groups	2
	x12 No. of modules	2
	x12-5 DMS module	2
	x12-8 Finance module	1
	x12-9 Accounting module	2
	x12-10 HRM module	1
	x12-12 SCM module	1
Situational	x16 No. of internal project members	10
	x17 No. of external consultants	10
	x21 Team maturity	1
	x31 Maturity of processes	3
	x33 Stability of organisation	7
	x35 Commitment management	3
	x36 Satisfaction with ERP system	2

Their different occurrence frequencies are clearly recognisable. The following five are the most frequented:

With a relationship to 11 of the cost fields, the most frequently influencing variable is “number of users”. This value is followed by the two cost drivers “number of external consultants” and “number of internal project managers”, each correlating with 10 cost fields.

With 9 connections, the cost drivers “number of locations” and “revenue” share the third place. “Number of departments or units” correlates with 8 of the cost drivers.

On the middle level, the two variables “stability of organisation” and “ERP system – SAP” show correlations to cost fields.
The cost driver “stability of organisation” features 7 negative correlations with the cost fields y1 to y5, y8 and y9. Showing correlations to 6 cost fields, “ERP system – SAP“ is the second most frequently arising correlation on the middle level.
All other cost drivers show between 1 and 4 relationships to the cost fields.

In total, 23 significant cost drivers could be identified. Based on this result, some of the discussed cost drivers can be rejected. The literature review discovered 35 potential cost drivers; this study is able to eliminate 20 of them as not significant.[17]
An overview of the irrelevant cost drivers is presented in the table below:

Table 6.45: Findings RQ2: Cost drivers without significant correlation

Organisational cost drivers	Technical cost drivers	Situational cost drivers
x3: No. of total employees	x7: No. of modifications	x18: Ratio external / internal
	x13: Complexity of data	x19: Fit of ERP system / organisation
	x14: Complexity of interfaces	x20: Team quality
	x15: Complexity of reports	x22: Team composition
		x23: Availability of management
		x24: Availability of business users
		x25: Consulting experience
		x26: Consulting quality
		x27: Critical attitude of users
		x28: User quality
		x29: Employee involvement
		x30: Management involvement
		x32: Complexity of business processes
		x34: Willingness to change
		x35: Commitment management

As approached in the sub-chapters, some of these results seem to be apparent whilst others do not feature a clear coherence. Moreover, some relationships between cost drivers and cost fields were expected, but did not show a significant relationship at all. These findings will be discussed in the following.

[17] The two variables “x5 ERP system” and “x12 type of modules” have dummy variables: x5 has 6 and x12 has 15. The total number of variables is 54.

Beginning with one of the anticipated relationships, the most frequently arising positive correlation "number of users" appears to make sense and has been validated in other studies before. In total, the variable has a connection to almost every cost field, with the exception of y10. Among them, the variable shows 2 high, 8 moderate and 1 low correlations.

It seems to be logical that these affect the amount of internal and external personnel costs during the three lifecycle phases.

During the evaluation phase, the ERP requirements of each user need to be gathered, defined and analysed by either internal or external personnel. The guideline for required costs could be that the higher the number of users, the more requirements need to be identified. During the implementation, every user is more or less involved in the project as a key user providing company-related know-how of business processes and needs to be trained at this stage of the lifecycle. This generates costs for internal personnel (y3) and also for external personnel (y4).

The positive connection between the number of users and the software costs during implementation as well as the hardware and licence costs during the implementation and evaluation phase appears to be plausible. However, it appears to be unapparent that there is no significant connection to y10 (software costs during maintenance).

The positive correlations between "number of external consultants" as well as "number of internal project members" and nearly all cost fields give rise to some questions. It appears to be obvious, of course, that the costs for external personnel increase with the number of external consultants and that the number of internal project members requires a corresponding expense for internal personnel.

But why should the software, licence and hardware costs in both the implementation and the evaluation phase generate more expenses with an increasing number of external consultants and of internal project members? That relationship seems to be unapparent.

Both of the variables point to the project size, but do not have a clear connection.

However, as mentioned earlier, the correlation analysis is not able to determine a causal connection between variables. All possible explanations provided are merely assumptions. This analysis is more about finding suitable indicators which influence the costs of ERP systems. These indicators can, of course, be unapparent, but the fact is that they have a relation to the dependent variable.

This leads over to the next frequently identified connections “number of locations”, “number of units or departments” and “revenue”.
For the two first named connections, the same conclusion applies as for the above-mentioned connections: some correlations appear to be obvious, whilst others remain unclear.
The correlations to software and hardware seem to be plausible. These costs increase with a rising number of locations and departments. The more software and hardware needs to be implemented and maintained, the higher the resulting costs.
In contrast, the correlations to the licence costs are more difficult to explain. Licence costs usually occur for the number of users and the number of modules. It is not that obvious why they would relate to the number of locations and the number of units or departments. They could be an indicator of the project size, like the next cost driver “revenue”.

The variable “revenue” correlates with 9 cost fields. It seems to have no apparent connection to any of them, but appears to be an important indicator for assessing the quantity of expenses within them. In 5 cost fields, it provides a moderate guide for assessing the cost fields. Four cost fields show at least a low correlation.

After regarding the most frequently arising correlations for their plausibility, the other remarkable findings are demonstrated below.
One of the most surprising results is the positive correlation with one of the famous ERP vendors, SAP. In total, the analysis identified 6 correlations; four show a moderate relationship and two a low relationship. Choosing this vendor seems to generate increased expenses within the cost fields y3, y4, y8, y9 and y10 compared to Microsoft, proALPHA, SAGE, and ABAS, which receive their own sub-variables. In this context, the negative correlation to the variable “ERP system – other” also needs to be mentioned.
Selecting one of the smaller providers appears to cause lower costs in at least four cost fields, which are internal and external personnel costs during the evaluation and implementation phase. To explain this would require analysing the different approaches of big vendors in contrast to the small ones, which cannot be done in this thesis.

The negative relationship of the variable “stability of organisation” is a very astonishing finding. That result means that the costs decrease with a higher organisational stability.

This effect is discovered within 7 cost fields. Two of these correlations are moderate, whilst the five remaining ones present a low connection.

One would have expected a positive relationship; these results are a bit paradox and difficult to interpret. It would have been more logical that stable organisations are willing to spend more on internal and external personnel during the three lifecycle phases as well as on software during the implementation phase.

The fact that the “soft skill” variables, like “team maturity”, “commitment management” and “satisfaction” only play a role during the evaluation phase is also an unexpected result. They appear to have no impact on the expenses during the implementation and maintenance phase. One would have expected them to have more influence; this is an astonishing result.

A higher influence was also expected for the technical variables “number of modules”, “number of interfaces” and “number of reports”. All of these variables show just low relationships to at least two cost fields. They do not seem to be that relevant for assessing the expenditure in ERP projects, which seems difficult to understand.

The variables presented in table 7.45 that show no significance at all also have some unpredicted outcomes. One would have anticipated them to influence the costs of ERP systems; it is unexpected having to dismiss the following variables as irrelevant in this thesis: “number of modifications”, “complexity of data”, “complexity of interfaces”, “complexity of reports”, “fit of ERP system / organisation”, “team quality”, “willingness to change” and “motivation of implementation team”. One must admit that this rejection is wondrous. But, as mentioned before, causality and reasons why some variables influence the cost fields and others have no effect cannot be provided in this correlation analysis. Further research must be conducted to understand these correlations.

Even if not all identified outcomes can be explained, the analysis of RQ2 managed to yield helpful results.

The 23 indicators identified might create added value for ERP costumers and project managers in helping them to have an overview of which cost drivers influence the costs at what stage of the lifecycle.

Many cost drivers are discussed within the literature; this study managed to exclude 20 of them. This may contribute to focussing on the relevant cost drivers and could increase the awareness of potentially emerging costs. This is important because vendor quotations do not picture all costs actually arising during the whole lifecycle, but are rather a first indication. Many companies had already experienced an unpleasant surprise because the project is not invoiced based on the quotation but on the actual efforts by the ERP provider. Consequently, a better knowledge about the relevant cost drivers helps to avoid cost overruns. Having identified some key influencing indicators, this study contributes to this aim and provides a guideline.

Furthermore, the identified indicators might be helpful not only for consumers but also for vendors. All vendors have their own systems for calculating ERP projects. These findings could at least help to review them.

Besides identifying the relevant cost drivers influencing each cost field, the results provide a first starting point for developing a cost estimation model. The 23 cost drivers determined to be relevant might enable the prediction of ERP costs. This approach will be specified in more detail in the next chapter.

6.4 RQ3: ANALYSIS & FINDINGS

This thesis has so far managed firstly to identify a percentage cost structure of cost fields for the costs arising during each ERP lifecycle phase and for its whole lifespan, and secondly verified 23 relevant cost drivers to have an impact on ERP costs.

How can these findings be used to predict ERP costs? This section would like to develop a formula for estimating them.

The findings allow for the development of two different approaches:

The first approach is to predict the costs for each single cost field. Each cost field will be regarded as a dependent variable (y1 – y12), and the relevant cost drivers of each cost field determined in RQ2 will be analysed.

The second approach is to predict the total costs of the ERP lifespan and to distribute these costs according to the average percentage cost structure identified in RQ1. The total costs are regarded as dependent variable y21 and its correlating cost drivers as independent variables. Both approaches apply multiple regression analysis. Their results are assessed by using the MMRE (Mean Magnitude of Relative Error).

$$MRE = \left| \frac{(Actual\ Effort - Estimated\ Effort)}{Actual\ Effort} \right| \times 100$$

$$MMRE = \frac{1}{n} \sum_{i=0}^{n} MRE_i$$

The MMRE decribes the mean error of the actual effort one must expect for the estimated effort in percentage form. Using MMRE overestimates as well as underestimates are considered.

Since multiple regressions are used in both approaches, this chapter firstly gives more details about this data analysis tool. Afterwards, it explains both approaches and shows their details. The outcome is discussed in the result comparison.

6.4.1 MULTIPLE REGRESSION ANALYSIS

Multiple regression is a very common method for analysing data with several independent variables and one dependent variable (Gliner et al., 2009). It is used for estimating the value of the dependent variable based on the values of two or more other independent variables (Burns & Burns, 2008). The linear combination of these independent variables is created by using a computed multiple regression, so that the "overall correlation R of the independent variables and the dependent variable is maximized, and the error in the prediction is minimized" (Gliner et al., 2009, p. 330). The regression quantifies the impact that each of the independent variables has on the dependent variable (Burns & Burns, 2008). The general equation is the following:

$$Y = b_0 + b_1X_1 + b_2X_2 + b_3X_3 + b_nX_n$$

b: unstandardised regression coefficient

X. independent variable

In this thesis, multiple regression analysis aims to identify a proper predictive ability of input variables to significantly forecast the ERP costs. It should be done with as small a set as possible in order to decrease complexity.
Since stepwise multiple regressions add and/or delete candidate predictors until one ‘best’ subset of predictor variables is achieved (Kuhn, 2003), it seems to be the most appropriate method. It yields results that show the best combination of possible independent variables for predicting the dependent variable. In other words, stepwise regression inserts or removes predictor variables systematically until inserting or removing further predictor variables in the regression equation yields no further improvement.

The chosen entity method is forward. The selection process from the candidates’ predictor pool starts with the independent variable featuring the highest correlation with the dependent variable (Backhaus et al., 2008). “If this process is not significant, the process stops […]. If this correlation is significant, the process goes on to the next step, adding the predictor variable which, in combination with this first one, has the highest multiple R. If it is a significant improvement, the [computer] programme repeats the process until either all predictor variables are included or adding any of the remaining ones fails to generate a significant improvement. This method allows judging the relative importance of each variable […]” (Burns & Burns, 2008, p. 397).

Stepwise regression requires two significance levels: one for adding variables and one for removing variables. The cut-off probability for adding variables should be less than the cut-off probability for removing variables so that the procedure does not get into an infinite loop[18]. In this thesis, the input criterion is $p > 0.05$ which is equivalent to a significance level of 95%. The cut-off value is $p > 0.10$.

Once the stepwise multiple regression has been conducted, the regression function and its individual repressor will be assessed.

[18] NCSS Statistical Software :
http://ncss.wpengine.netdna-cdn.com/wp-content/themes/ncss/pdf/Procedures/NCSS/Stepwise_Regression.pdf

6.4.2 APPROACH I: COST PREDICTION OF EACH COST FIELD

This approach predicts the costs of each cost field, for one thing. For another thing, it allows regarding the total costs by adding up the cost fields.

Each cost field (y1-y12) is predicted by using the correlating independent variables (cost drivers) identified in RQ2. These findings are used as input variables for developing the formula.
Stepwise multiple regression adds or deletes them as long as the best selection is yielded. This procedure is conducted for each of the twelve cost fields.

For each of them, the table below gives a holistic overview about the input variables, the deleted cost driver candidates, the resulting formula, and the best selection of input variables.

Table 6.46: Approach I: Overview of the subset of significant variables and prediction formula

Cost Field	**Input Variables**	**Deleted Input Variables**	**Best Subset**	**Prediction formula**	Qty. of remaining input variables
y1 Internal Personnel Evaluation	x1: No. of locations x10: No. of total users x31: Maturity of processes x33: Stability of organisation x35: Commitment management x36: ERP satisfaction	x10: No. of total users x31: Maturity of processes x33: Stability of organisation x35: Commitment management x36: ERP satisfaction	x1: No. of locations	y1 = 16,787 * x1	1
y2 External Personnel Evaluation	x1: No. of locations x10: No. of total users x21: Team maturity x31: Maturity of processes x33: Stability of organisation x35: Commitment management x36: ERP satisfaction	x10: No. of total users x21: Team maturity x31: Maturity of processes x33: Stability of organisation	x1: No. of locations x35: Commitment management	y2 = 18,883 * x1 – 4,553 * x35	2
y3 Internal Personnel Implementation	x1: No. of locations x2: No. of organisational units or depts. x4: Revenue x5-1: System - SAP x5-6: System - other x10: No. of total users x16: No. of internal project members x17: No. of external consultants x33: Stability of organisation	x1: No. of locations x2: No. of organisational units or depts. x4: Revenue x5-6: System - other x16: No. of internal project members x33: Stability of organisation	x5-1: System - SAP x10: No. of total users x17: No. of external consultants	y3 = 1,011 * x10 + 9,940 * x17 + 97,909 * x5-1	3

y4 External Personnel Implementation	x1: No. of locations x2: No. of organisational units or depts. x4: Revenue x5-1: System - SAP x5-6: System - other x10: No. of total users x12-9: Accounting module x12-10: HRM module x16: No. of internal project members x17: No. of external consultants x33: Stability of organisation	x1: No. of locations x4: Revenue x5-1: System - SAP x12-10: HRM module x16: No. of internal project members x33: Stability of organisation	x2: No. of organisational units or depts. x5-6: System - other x10: No. of total users x12-9: Accounting module x17: No. of external consultants	y4 = (3,923 * x2) – (66,109 * x5-6) + (286 * x10) + (51,612* x12-9) + (17,499 * x17)	5
y5 ERP Software Implementation	x1: No. of locations x2: No. of organisational units or depts. x4: Revenue x10: No. of total users x11: No. of user groups x16: No. of internal project members x17: No. of external consultants x33: Stability of organisation	x1: No. of locations x2: No. of organisational units or depts. x4: Revenue x16: No. of internal project members x17: No. of external consultants x33: Stability of organisation	x10: No. of total users x11: No. of user groups	y5 = 870 * x10 + 3497* x11	2
y6 Licence Implementation	x1: No. of locations x2: No. of organisational units or depts. x4: Revenue x10: No. of total users x12-8: Finance module x16: No. of internal project members x17: No. of external consultants x31: Maturity of processes	x2: No. of organisational units or depts. x4: Revenue x10: No. of total users x12-8: Finance module x16: No. of internal project members x31: Maturity of processes	x1: No. of locations x17: No. of external consultants	y6 = 14,015 * x17 + 23,787 * x1	2

y7 Hardware Implementation	x2: No. of organisational units or depts. x4: Revenue x5-1: ERP system - SAP x10: No. of total users x16: No. of internal project members x17: No. of external consultants	x4: Revenue x10: No. of total users x16: No. of internal project members x17: No. of external consultants	x2: No. of organisational units or depts. x5-1: ERP system - SAP	y7 = 5,630 * x2 + 70,212 * x5-1	2
y8 Internal Personnel Maintenance	x1: No. of locations x2: No. of organisational units or depts. x4: Revenue x5-1: ERP system - SAP x5-6: ERP system - other x6: No. of interfaces x8: No. of reports x9: No. of EDIs x10: No. of total users x16: No. of internal project members x17: No. of external consultants x33: Stability of organisation	x4: Revenue x5-1: ERP system - SAP x5-6: ERP system - other x6: No. of interfaces x9: No. of EDIs x17: No. of external consultants x33: Stability of organisation	x1: No. of locations x2: No. of organisational units or depts. x8: No. of reports x10: No. of total users x16: No. of internal project members	y8= 704 * x10 + 257 * x8 – 13,808 * x1 + 2,712 * x2 – 2,747 * x16	5
y9 External Personnel Maintenance	x1: No. of locations x2: No. of organisational units or depts. x4: Revenue x5-1: ERP system - SAP x5-6: ERP system - other x10: No. of total users x12: No. of modules x12-5: DMS module x16: No. of internal project members	x1: No. of locations x2: No. of organisational units or depts. x4: Revenue x5-1: ERP system - SAP x5-6: ERP system - other x12: No. of modules x12-5: DMS module x16: No. of internal project members x33: Stability of organisation	x10: No. of total users 17: No. of external consultants	y9=141 * x10 + 2,795 * x17	2

	x17: No. of external consultants 33: Stability of organisation x35: Commitment management	x35: Commitment management			
y10 ERP Software Maintenance	x5-1: ERP system - SAP x16: No. of internal project members x17: No. of external consultants	x5-1: ERP system - SAP x16: No. of internal project members	x17: No. of external consultants	y10= 2024 * x17	3
y11 Licence Maintenance	x1: No. of locations x2: No. of org. units or depts. x4: Revenue x9: No. of EDIs x10: No. of total users x11: No. of user groups x12: No. of modules x12-5: DMS module x12-9: Accounting module x16: No. of internal project members x17: No. of external consultants	x1: No. of locations x2: No. of org. units or depts. x4: Revenue x9: No. of EDIs x12: No. of modules x12-9: Accounting module x16: No. of internal project members	x10: No. of total users x11: No. of user groups x12-5: DMS module x17: No. of external consultants	y11= 1,866 * x17 + 98 * x10 + 416 * x11 + 9,056 * x12-5	4
y12 Hardware Maintenance	x4: Revenue x10: No. of total users x16: No. of internal project members x17: No. of external consultants	x4: Revenue x16: No. of internal project members	x10: No. of total users 17: No. of external consultants	y12= 58 * x10 + 893 * x17	2

Table 6.46 provides an overview of the resulting formula and the best subset of variables to predict the costs of each cost field y1 – y12.

In general, it is surprising that more than half of the input variables are deleted from the models and reduced to a very small number in each cost field. This seems to be rather contradictory since ERP costs are complex and expected to be dependent on lots of different cost drivers. The two cost fields with the highest number of subset variables are y4 and y8. Both generated formulas include five cost drivers. With only one subset variable, the cost field with the smallest number of subset variables is y1.

Furthermore, the exclusion of some variables is astonishing. The most surprising one is the exclusion of the variable "number of users" in y6 (licence cost implementation). Usually, licence is invoiced per user and consequently is bound to have a strong dependence. An explanation could be that licence models vary strongly between different vendors. Some vendors do not have licence costs at all but charge maintenance costs for their software products. However, in this context, it is remarkable that the variable is included in y10 (licence costs maintenance) into the resulting formula. This might be associated with the fact that the implementation phase ends before the ERP system goes live. During this "testing stage", it is unlikely that all future users are being equipped with a licence because of the maintenance costs that would then result for that licence. One can expect that this is done upon or shortly before the transition to the go-live.
The variable "number of users" was also expected to have a much stronger influence in y1 and y2.

Furthermore, the exclusion of the variable "maturity of processes" in y1 and y2, "number of interfaces" in y8, and "number of modules" in y9 and y11 is unforeseen. These variables were expected to have a much stronger influence.

In order to evaluate the formula, the percentage errors showing the average deviation between the estimated value and the actual value of each cost field need to be studied. This was done with the help of the MMRE described in chapter 6.4.

The table below gives an overview of the results calculated for each cost field.

Table 6.47: MMRE of each cost field

Y1	Y2	Y3	Y4	Y5	Y6	Y7	Y8	Y9	Y10	Y11	Y12	∑(y1 to y12)
±175	±345	±197	±301	±121	±245	±194	±791	±183	±112	±54	±103	±49

Table 6.47 shows a divergent picture of the MMRE of each cost field. Whilst some cost fields show “just” about +/- 54%, others present very high values in the three-digit level. Calculating the total costs of the ERP lifespan by adding up the twelve cost fields presents a relatively “good” result. A percentage estimation error of +/- 49% is achieved. Apparently, the percentage errors neutralise each other.

Despite this comparatively “good” result for the total ERP costs, the prediction of each cost field remains unsatisfying. One reason for this could be the existence of some strong influencing outliers. The data was analysed for them and adjusted by 10% of the strongest outlier results. Table 6.48 shows that this results in a much better outcome:

Table 6.48: MMRE adjusted by 10% outliers

Y1	Y2	Y3	Y4	Y5	Y6	Y7	Y8	Y9	Y10	Y11	Y12	∑(y1 to y12)
±87	±232	±79	±123	±61	±95	±76	±292	±54	±55	±29	±64	±37

Through the exclusion of the outliers, the prediction result of each cost field could be improved substantially. The MMRE of each cost field estimation could be reduced about 2/3 of its previous value. Also, the value of the total lifecycle costs (by adding up absolute errors of the cost fields y1 to y12) could be improved from +/- 54% to +/- 37%.

Nevertheless, the evaluation shows that cost fields are hard to predict with this approach. The MMRE shows a high average deviation between the estimated value and the actual value, except for y11. This results in the disappointing realisation that cost fields cannot be predicted accurately with the developed formula.

However, this approach provides a relatively good formula for predicting the total ERP costs. Predicting each field with the resulting formulas presented in table 6.46 provides a result which averagely deviates about +/- 37%. Compared to the MMRE results of previous empirical research, this value is close to the 40% achieved by Widmer (2004) but even shows a slightly better accuracy of estimate.

6.4.3 APPROACH II: DISTRIBUTING PREDICTED TOTAL COSTS ACCORDING TO IDENTIFIED COST STRUCTURE IN RQ1

Instead of estimating each cost type separately, this second approach predicts the total ERP costs and distributes them according to the average percentage cost structure identified in RQ1. The total costs are referred to as dependent variable y21.

Since a correlation analysis was not conducted in RQ2 for the whole ERP costs, this chapter has to start with finding the relevant cost driver candidates. This is done by correlation analysis.

The correlation analysis identified 13 cost drivers with high, moderate and low relationships. The highest substantial relationships are discovered between the total ERP costs and the cost drivers “no. of total users” and “no. of external consultants”.
Moderate relationships to the costs are found for the variables “no. of locations”, “no. of org. units or depts.”, “revenue”, “no. of internal project members” and “stability of organisation”.
The variables identified as showing only a low correlation are “ERP system – other”, “no. of modules”, “DMS module”, “Finance module”, “maturity of processes”, and “commitment management”.

Table 6.49 gives an overview of the correlating cost drivers arranged in a descending order and sorted by their correlation coefficient from +1 to -1.

Table 6.49: Correlation between total lifecycle costs and cost drivers

Independent Variables: Cost Drivers	N	Pearson Corr.	r^2	Signif. 2-tailed (p-Value)	Significance level	Interpretation
x10 No. of total users	72	0.869	76%	0	**	high correlation
x17 No. of external consultants	70	0.744	55%	0	**	high correlation
x2 No. of org. units or depts.	72	0.602	36%	0	**	moderate correlation

x1 No. of locations	72	0.576	33%	0	**	moderate correlation
x4 Revenue	69	0.576	33%	0	**	moderate correlation
x16 No. of internal project members	70	0.504	25%	0	**	moderate correlation
x5-1 ERP system - SAP	72	0.483	23%	0	**	moderate correlation
x12-5 DMS module	72	0.327	11%	0.005	**	low correlation
x12-8 Finance module	72	0.319	10%	0.006	**	low correlation
x12 No. of modules	72	0.317	10%	0.007	**	low correlation
x8 No. of reports	67	0.303	9%	0.013	*	low correlation
x13 Complexity of data	72	0.296	9%	0.011	*	low correlation
x12-12 SCM module	72	0.276	8%	0.019	*	low correlation
x9 No. of EDIs	70	0.262	7%	0.029	*	low correlation
x6 No. of interfaces	69	0.255	7%	0.034		low correlation
x12-9 Accounting module	72	0.249	6%	0.035	*	low correlation
x12-3 Calculation module	72	0.237	6%	0.045	*	low correlation
x12-10 HRM module	72	0.224	5%	0.059		low correlation
x12-4 PM module	72	0.184	3%	0.123		slight correlation
x12-11 Production module	72	0.153	2%	0.2		slight correlation
x14 Complexity of interfaces	70	0.139	2%	0.252		slight correlation
x11 No. of user groups	71	0.132	2%	0.274		slight correlation
x5-2 ERP system - Microsoft	72	0.115	1%	0.336		slight correlation
x18 Ratio internal/external	69	0.097	1%	0.425		slight correlation
x12-2 MM module	72	0.086	1%	0.473		no correlation
x21 Team maturity	68	0.067	0%	0.587		no correlation
x5-3 ERP system - proALPHA	72	0.052	0%	0.667		no correlation
x15 Complexity of reports	66	0.048	0%	0.702		no correlation
x12-1 Purchasing module	72	0.044	0%	0.714		no correlation
x12-6 Sales module	72	0.042	0%	0.726		no correlation
x20 Team quality	71	0.028	0%	0.818		no correlation
x32 Complexity of bus. processes	72	0.022	0%	0.831		no correlation
x12-7 CRM module	72	-0.001	0%	0.993		no correlation
x29 Employee involvement	70	-0.004	0%	0.976		no correlation
x7 No. of modifications	68	-0.01	0%	0.935		no correlation
x23 Availability of management	69	-0.014	0%	0.911		no correlation
x28 User quality	68	-0.034	0%	0.782		no correlation
x36 Satisfaction with ERP system	71	-0.044	0%	0.716		no correlation
x5-5 ERP system - ABAS	72	-0.053	0%	0.657		no correlation
x22 Team composition	70	-0.08	1%	0.513		no correlation

x26 Consulting quality	68	-0.081	1%	0.51		no correlation
x19 Fit of system/organisation	71	-0.096	1%	0.428		no correlation
x34 Willingness to change	70	-0.107	1%	0.376		slight correlation
x25 Consulting experience	67	-0.108	1%	0.386		slight correlation
x27 Critical attitude of users	70	-0.109	1%	0.371		slight correlation
x5-4 ERP system - SAGE	72	-0.118	1%	0.322		slight correlation
x12-15 MDA module	72	-0.126	2%	0.291		slight correlation
x12-14 PDA module	72	-0.128	2%	0.285		slight correlation
x30 Management involvement	70	-0.184	3%	0.127		slight correlation
x24 Availability of business users	68	-0.188	4%	0.126		slight correlation
x12-13 Detailed planning module	72	-0.192	4%	0.106		slight correlation
x31 Maturity of processes	69	-0.311	10%	0.009	**	low correlation
x35 Commitment management	71	-0.322	10%	0.006	**	low correlation
x5-6 ERP system - other	72	-0.38	14%	0.001	**	low correlation
x33 Stability of organisation	70	-0.444	20%	0	**	moderate correlation

So far, this chapter identified 13 relevant relationships between the cost drivers and the costs for the whole ERP lifecycle. These are summarised in the next table:

Table 6.50: Findings total lifecycle costs (y21)

Cost Driver Variables	Correlation (r)	r^2	Correlation strength
Positive Correlations			
x10 No. of total users	+0.869	76%	high correlation
x17 No. of external consultants	+0.744	55%	high correlation
x_{2} No. of org. units or depts.	+0.602	36%	moderate correlation
x_{1} No. of locations	+0.576	33%	moderate correlation
x_{4} Revenue	+0.576	33%	moderate correlation
x_{16} No. of internal project members	+0.504	25%	moderate correlation
$x_{5\text{-}1}$ ERP system - SAP	+0.483	23%	moderate correlation
$x_{12\text{-}5}$ DMS module	+0.327	11%	low correlation
$x_{12\text{-}8}$ Finance module	+0.319	10%	low correlation
x_{12} No. of modules	+0.317	10%	low correlation
Negative Correlations			
x_{35} Commitment management	-0.322	10%	low correlation
$x_{5\text{-}6}$ ERP system - other	-0.380	14%	low correlation
x_{33} Stability of organisation	-0.444	20%	moderate correlation

These 13 identified cost drivers are the input variables for the stepwise regression model. As mentioned earlier, the stepwise regression adds and deletes cost driver candidates until the best model selection is yielded.

Table 6.51 provides an overview of the resulting formula and the best subset of variables to predict the costs of the total lifecycle (y21).

Table 6.51: Approach II: Overview of the subset of significant variables and prediction formula for total lifecycle costs (y21)

Input Variables	Deleted Input Variables	Best Subset	Prediction Formula	Number of Variables
x1 No. of locations	x1 No. of locations	x2 No. of org. units or depts.	y21 = 63,905 * x2 + 9,388 * x10 + 82,041 * x12 + 161,572 * x17 – 116,052 * x35	5
x2 No. of org. units or depts.	x4x4 Revenue	x10: No. of total users		
x4 Revenue	x5-1 ERP system - SAP	x12 No. of modules		
x5-1 ERP system - SAP	x5-6 ERP system - other	x17 No. of external consultants		
x5-6 ERP system - other	x12-5 DMS module	x35: Commitment management		
x10: No. of total users	x12-8 Finance module			
x12 No. of modules	x16 No. of internal project members			
x12-5 DMS module	x33: Stability of organisation			
x12-8 Finance module	x35: Commitment management			
x16 No. of internal project members				
x17 No. of external consultants				
x33: Stability of organisation				
x35: Commitment management				

Table 6.51 gave an overview of the input variables, the excluded candidates, the resulting formula and the best subset.
The exclusion of the candidates “internal project member” and “number of locations” is surprising. These variables were expected to influence the total ERP costs a lot more.
An impressive result is the strong influence of the variable “commitment management”. Being one of the five variables in the model, the motivation and dedication of employees seem to be very important for the success of ERP projects.

The total ERP costs should now be distributed according to the cost structure identified in RQ1. Table 6.52 repeats the findings:

Table 6.52: Percentage cost structure identified in RQ1

Cost fields	Average percentage quantity of total costs
y1 Internal Personnel Costs Evaluation	3.7%
y2 External Personnel Costs Evaluation	2.1%
y3 Internal Personnel Costs Implementation	9.8%
y4 External Personnel Costs Implementation	7.0%
y5 ERP Software Costs Implementation	12.5%
y6 Licence costs Implementation	7.0%
y7 Hardware Costs Implementation	4.8%
y8 Internal Personnel Costs Maintenance	14.8%
y9 External Personnel Costs Maintenance	10.9%
y10 ERP Software Costs Maintenance	5.3%
y11 Licence Costs Maintenance	16.3%
y12 Hardware Costs Maintenance	5.9%

Again, the result of the distribution should be assessed by regarding the quality criterion MMRE. This is shown in table 6.53.

Table 6.53: MMRE of each cost field

Y1	Y2	Y3	Y4	Y5	Y6	Y7	Y8	Y9	Y10	Y11	Y12	∑(y1 to y12)
±267	±288	±247	±196	±229	±169	±166	±467	±130	±84	±70	±105	±49

Table 6.54 presents a picture equal to the results in approach one. Since some of the cost fields have a very high MMRE, regarding the individual cost fields remains rather unsatisfying. One cost field deviates about +/- 70% while others present very high percentage errors in the three-digit range.

As in approach I, the calculation of the total costs of the ERP lifespan fields yields a rather good result. With a percentage estimation error of +/- 49%, the same result as in approach one is achieved.

At this stage, again, the data should be examined for and adjusted by the outliers. Adjusting the data by 10% of the strongest outliers results in a much better outcome.

Table 6.54 shows the results:

Table 6.54: MMRE of each cost field adjusted by 10% outliers

Y1	Y2	Y3	Y4	Y5	Y6	Y7	Y8	Y9	Y10	Y11	Y12	∑(y1 to y12)
±129	±182	±89	±67	±85	±84	±107	±206	±61	±64	±47	±68	±32

This result also does not provide satisfying yield of the cost field. Regarding their MMRE, none of the dependent variables show a reliable accuracy of estimate.

However, regarding the total costs, this approach is able to create a better result than approach I. Predicting each field by the resulting formulas presented in table 7.55 provides an outcome which averagely deviates about +/- 32%. Compared to the value of +/- 37% achieved in approach I, this is an improvement.

Both results are compared in more detail in the next chapter.

6.4.4 RESULT COMPARISON

Each of the two chapters above presented one approach to estimate ERP costs. This chapter will compare the results in more detail.

Table 6.55 shows the percentage error of each approach adjusted by 10% outliers.

Table 6.55: Comparison MMRE of approach I and approach II

Approach	Y1	Y2	Y3	Y4	Y5	Y6	Y7	Y8	Y9	Y10	Y11	Y12	∑(y1 to y12)
I	±87	±232	±79	±123	±61	±95	±76	±292	±54	±55	±29	±64	±37
II	±129	±182	±89	±67	±85	±84	±107	±206	±61	±64	±47	±68	±32

As is shown in the table above, both approaches produce a high MMRE for the single cost fields, and the result is rather disappointing. Values above +/- 50% are not accurate enough to make reliable predictions of the cost fields. Just y11 in Approach I yields a satisfying result with a value of +/- 29%.
When comparing the results of both approaches, one might say that Approach I produces better predictions for estimating the costs of the cost fields. Approach I seems to be more appropriate as it delivers better outcomes in eight of the twelve cost fields.
Nevertheless, the result for predicting cost fields is not satisfying.

Fortunately, the percentage errors of each cost field appear to neutralise each other, resulting in good values to estimate the total costs. One reason for that might be deferrals between cost fields and the lifecycle phases.
When comparing the achieved MMRE of 32% to previous empirical studies, one identifies this regression model to be the best. It is 8% more accurate than the one suggested by Widmer in 2004.
The table below restates the MMRE of previous empirical studies and includes the results of this study.

Table 6.56: Achieved MMRE compared to other empirical studies

Method	**Author**	**MMRE in %**
Social Choice (Copeland)	(Koch & Mitlöhner, 2010)	543
Social Choice (Borda)	(Koch & Mitlöhner, 2010)	958
Multiple Regression Analysis	(Koch & Mitlöhner, 2010)	1159
	(Widmer, 2004)	40
	(Myrtveit & Stensrud, 1999)	127
	This Study	**32**
Case-based Reasoning (ANGEL)	(Koch & Mitlöhner, 2010)	48
	(Myrtveit & Stensrud, 1999)	154
DEA	(Koch & Mitlöhner, 2010)	155

This result must be evaluated as a very good result. It seems to achieve a better accuracy of estimation than all approaches ever stated before.

In conclusion, the second approach, with a value of +/- 32% , makes more precise predictions than approach I for predicting total lifecycle costs and seems to achieve better results than the previously developed approaches. This is done by the following regression model.

*Total costs = 63,905 * x2 + 9,388 * x10 + 82,041 * x12 + 161,572 * x17 – 116,052 * x35*

x2: No. of org. units or depts.
x10: No. of total users
x12: No. of modules[19]
x17: No. of external consultants
x35: Commitment management[20]

[19] From: Purchasing, Sales, Material Management, CRM, SCM, Finance, Project Management, DMS, HR, Calculation, Production Planning, Detailed Production Planning, Capture of Production Data, Machine Data Collection, Personnel Accounting.

[20] Assessment between 1 to 10

In order to examine its quality, the formula should be assessed for its fitness to reality. This is done in the next chapter.

6.4.5 FORMULA ASSESSMENT

Approach II was calculated by applying stepwise multiple regression; it yields the following regression model:

*Y21 = 63,905 * x2 + 9,388 * x10 + 82,041 * x12 + 161,572 * x17 – 116,052 * x35*

To assess the quality of the formula, the following quality criteria need to be examined: firstly, the regression function, secondly the regressors, and thirdly the model premises. Each criterion receives its own sub-chapter in order to create a neat structure.

6.4.5.1 ASSESSEMT OF REGRESSION FUNCTION

In order to investigate if the regression function describes the reality of ERP costs, the quality of the regression function will be examined. This is done by applying the three quality criteria, coefficient of multiple determination (R^2), and F-statistics.

COEFFICIENT OF MULTIPLE DETERMINATION (R^2)

"R^2 describes the degree to which the predictor variables as a whole account for variations in the dependent variable" (Burns & Burns, 2008, p. 386). In this case, R^2 describes how many variations of y21 are explained with the variables x2, x10, x12, x17, and x35.

Since R^2 will be influenced by the number of regressors, the so-called adjusted R^2 which decreases the R^2 by a correcting quantity will be considered (Backhaus et al., 2008). Table 6.57 presents the results of the calculation of R, R^2 and the adjusted R^2.

Table 6.57.: R, R^2 and adjusted R^2

R	R^2	Adjusted R^2
.963	.928	.922

The adjusted R^2 of .922 shows that the regressors selected in approach II explain 92.2% of the variations. Only 7.8 % remain unaccounted for.

This can be assessed as a good value, expressing that the multiple linear regression functions fit quite well to the empirical data.

F-STATISTICS

Even if a regression function features a high "goodness-of-fit", this does not necessarily imply that the validity of the model is given if the regression function was estimated on too small a number of observations (Backhaus et al., 2008).

In order to verify if the high goodness-of-fit is a random result or based on the relationship of the data, a so-called hypothesis H_0 (null hypothesis) needs to be defined. The null hypothesis means that no relationship exists between the dependent variable (y21) and the independent variables x2, x10, x12, x17, and x35; therefore the standardised regression coefficients (ßx) are 0:

$$H_0 : ß2 = ß10 = ß12 = ß17 = ß35 = 0$$

The so-called F-test checks if H_0 can be rejected, which means that a relationship exists. Basically, the F-test consists of the calculation of an empirical F-statistic value (F_{emp}) which is then compared with a critical F-value (F_{crit}) which should not be underrun by the F_{emp}. Besides the significance level, this F_{crit} is based on the number of regressors and the number of observations.

For this regression model, this means that the value of F_{crit} should not be smaller than 2.37[21]. SPSS calculated an F_{emp} of 149.63, which is considerably higher than the F_{crit}.

The following applies: $F_{emp} > F_{crit}$ → H_0 is rejected

This leads to the result that the relationships of the variables within the regression model are significant (p-value =.000).

[21] Based on a significance level of 0.95, five dependent variables and 60 observations.

6.4.5.2 ASSESSMENT OF REGRESSORS

In the previous subsection, it was examined how the regression function as a whole is able to predict the total lifecycle costs (y21). In addition, it is necessary to verify the individual regressors of the regression function. On the one hand it will be examined which regressors have the strongest influence on the dependent variable, and on the other hand it will be examined if the regressors feature an appropriate significance level based on the t-statistics.

Table 6.58 presents the unstandardised regression coefficients (B) as well as the standardised beta coefficients (ß). The beta coefficients enable the comparability of the individual regressors within the regression function. They show that the total number of users (x10) has the strongest influence within the regression function. In contrast, the number of organisational units or departments (x2) carries the lowest weight within the regression model.

Table 6.58: Coefficient details of final regression model

Regressors	Unstandardised coefficients B	Standardised coefficients Beta	t_{emp}	Signif.
x10 No. of total users	9388.026	.467	6.855	.000
x17 No. of external consultants	161572.287	.349	5.305	.000
x2 No. of organisational units or depts.	63904.819	.289	4.563	.000
x35 Motivation of implementation team	-116051.463	-.407	-3.374	.001
x12 No. of modules	82041.165	.304	2.345	.022

The t-test will be conducted to verify the individual regressors. Here again, a null hypothesis H_0 will be defined for every individual standardised regression coefficient within the regression function:

$H_0 : ß2 = 0$

$H_0 : ß10 = 0$

$H_0 : ß12 = 0$

$H_0 : ß17 = 0$

$H_0 : ß35 = 0$

Then, just like in the F-test, calculation of an empirical t-statistic value (t_{emp}) will be compared with a critical t-value (t_{crit}), which should not be underrun by the t_{emp}. With a significance level of 0.95, a t_{crit} of 2,000 should not be underrun from each individual absolute t_{emp} value of the regressors. The results of the t_{emp} are also presented in table 7.59. It can be seen that no absolute t_{emp} value falls below the t_{crit}. This entailed that the null hypothesis had to be rejected for all regressors within the regression functions. The influence of all regressors is thus significant.

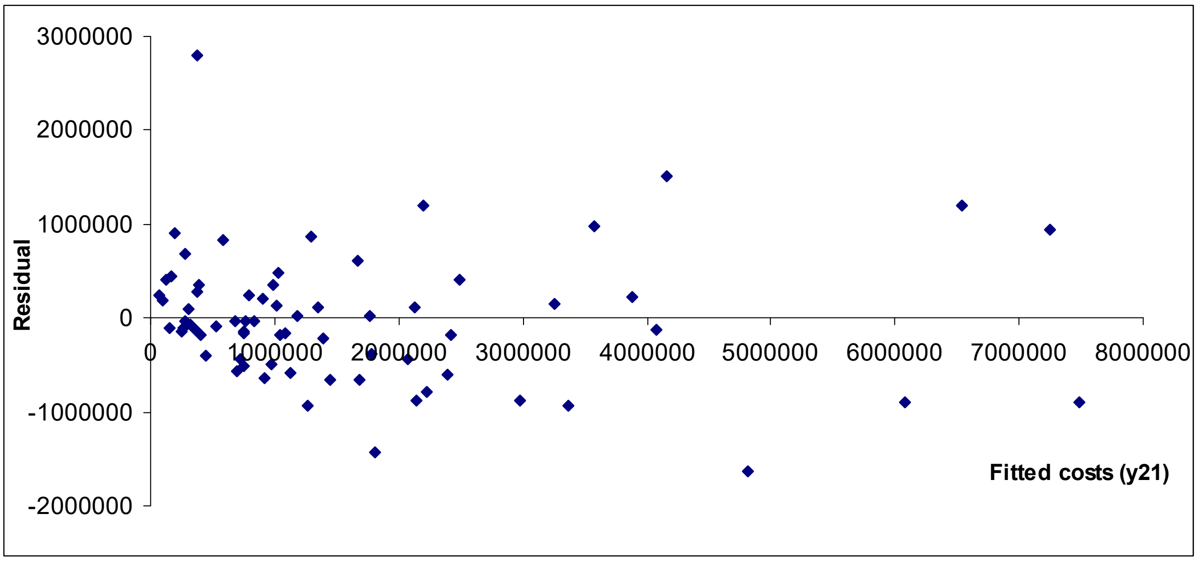
6.4.5.3 ASSESSMENT OF MODEL PREMISES

Finally, it is necessary to assess whether the premises for the development of a regression model are fulfilled. Following the assessment premises proposed by Backhaus et al. (2008), the following premises were evaluated.

HOMOSCEDASTICITY

The first model premise to be assessed is the existence of homoscedasticity. It describes whether the residuals of the dependent variable feature similar variances. The absence of these similar variances would indicate heteroscedasticity. However, in order to analyse if the data are homoscedastic, the variances of the residuals should be visually inspected. Figure 6.6 shows the residuals resulting from the differences between actual and fitted costs of y21.

Figure 6.6: Residuals resulting from the difference between actual and fitted costs



If heteroscedasticity exists, "there is a clear systematic pattern and from this we could infer that the residual variance in not constant" (Baddeley & Barrowclough, 2009, p. 145). The scatter-plot of the residuals often shows a cognisable relationship by forming a triquetrous pointed pattern (Backhaus et al., 2008).

However, in this case, no such systematic pattern is evident. Therefore, there is no indication for the existence of heteroscedasticity.
At this stage, it must be noted that one reason for the non-existence of heteroscedasticity may be the relatively low response rate of 4.5%.
However, the first model premises can thus be assessed as fulfilled.

NON-EXISTENCE OF AUTOCORRELATION
Another model premise is that the residuals must be free from autocorrelation. "Autocorrelation (in the residuals) is a problem that occurs when residuals from a sample regression are not random because they correlate with each other over time or space" (Baddeley & Barrowclough, 2009, p. 165). Autocorrelation is a "violation of the assumption that the covariance between error terms is zero" (Baddeley & Barrowclough, 2009, p. 165).

Similar to the detection of heteroscedasticity, a visual inspection of systematic patterns of the residuals plots is recommended (Backhaus et al., 2008; Baddeley & Barrowclough, 2009). As has already been noted above, no systematic residual pattern could be determined.
Aside from visual detection, there are several mathematical diagnostic approaches, such as the Durbin-Watson test, quantifying systematic residual patterns (Baddeley & Barrowclough, 2009). The result of the Durbin-Watson test is a value between 0 and 4. Values close to 0 indicate a strong positive autocorrelation, while values close to 4 indicate a strong negative autocorrelation. Values lower than 1 or higher than 3 indicate autocorrelation. Values between 1.5 and 2.5 can be interpreted as acceptable, and values somewhere around 2 indicate that there is no autocorrelation. The result of the Durbin-Watson test for this regression analysis is 1.797 which indicates that the residuals from the sample regression are random.

Both the visual inspection and the Durbin-Watson test indicate no existence of autocorrelation; thus this premise of the regression model is fulfilled as well.

NON-EXISTENCE OF MULTI-COLLINEARITY

Another premise for a multiple linear regression model is that a linear relationship between the regressors of the regression model is not manifested (Backhaus et al., 2008). First evidence can be found by inspecting a correlation matrix of the dependent variables. Table 6.59 summarises the results again by only considering the dependent variables included in the regression model.

Table 6.59: Correlation matrix of regressors

	x2	x10	x12	x17	x35
x2	1.000				
x10	0.526	1.000			
x12	0.032	0.336	1.000		
x17	0.322	0.634	0.245	1.000	
x35	-0.006	-0.222	0.046	-0.329	1.000

Indeed, a moderate correlation exists between the variables x10 and x2 as well as between the variables x10 and x17, but is not as strong as a high or very high relationship that is close to r=1.000 (perfect correlation).

However, correlation coefficients only measure the relationship between two variables. For this reason, even when the correlation matrix does not show (very) high correlations, it is possible that multi-collinearity can exist (Backhaus et al., 2008). This can happen when one of the regressors features a relationship which can be explained by two or more other regressors within the regression model. Therefore it is proposed to analyse for every regressor a regression analysis using the other regressors as independent variables (Backhaus et al., 2008). The resulting R^2 for each regression is then used in the so-called 'tolerance' measure or 'VIF' (variance inflation factor) (Backhaus et al., 2008). Table 6.60 presents the results of the analysis of tolerance and VIF for the regression model.

Table 6.60 Collinearity statistics

Variables	Tolerance	VIF
x10: No. of total users	.267	3.748
x17: No. of external consultants	.286	3.491
x2: No. of organisational units or depts.	.310	3.225
x35: Motivation of implementation team	.085	11.734
x12: No. of modules	.074	13.562

There is no generally accepted definition of a critical value showing the existence of multi-collinearity. However, it is often suggested to investigate tolerance values lower than 0.1 or VIF higher than 10 more closely.

Indeed, the variables x35 and x12 show values closely under/over the suggested values, but it is not a logical explanation that these two variables can be totally explained by the other variables. It seems to be logical that the extent of an ERP implementation expressed by the variable x12 is in some respect dependent on the company size, which surely will influence the variable x10 and x2. These variables cannot entirely and precisely explain x12 on a logical basis.

For the variable x35, there is no logical reason why this variable could be explained by the other variables.

However, it was tested how the standard error of estimation would change when excluding the variables x12 and x35 from the regression model. After elimination of the two independent variables, the standard error of estimation then actually declined somewhat.

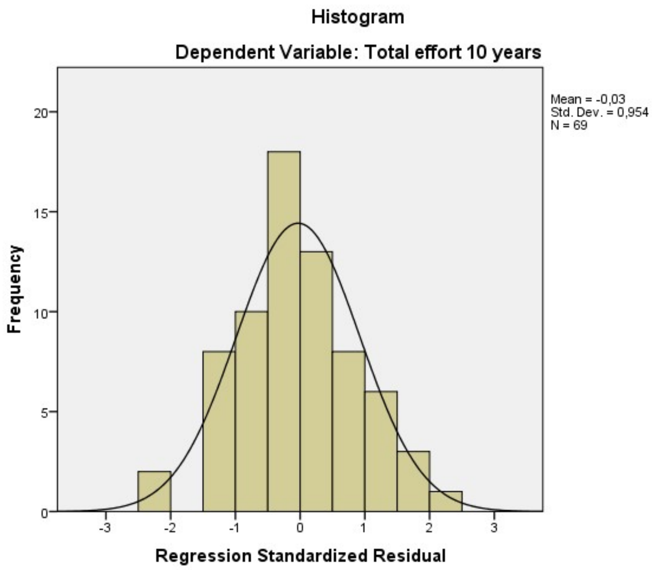
Therefore, it was decided to leave the variables x12 and x35 in the regression model even if there is at least a speculation about the existence of multi-collinearity with these variables. The premise of the non-existence of multi-collinearity is thus assessed as fulfilled, with reservations.

NORMALITY OF ERRORS

Finally, the assumption that the residuals are normally distributed should be assessed. The assumption of normal distribution of the residuals is especially important for significance testing (F-test, t-test) when the number of observations is smaller than 40 (Backhaus et al., 2008). Since there are 69 observations in this study, the normality of errors would not have any influence on the significance tests.

However, the histogram presented in figure 6.7 shows that there is a fairly normal distribution of the residuals.

Figure 6.7: Histogram of standardised residuals



The premise of the assumption of normal distribution thus counts as fulfilled.

6.4.6 SUMMARY & CONCLUSION RQ3

This chapter aimed to find a regression model for predicting, in a first step, each cost field and, secondly, the total costs of the ERP lifecycle.

Two different approaches were applied. The first approach predicted the costs for each single cost field. The second approach estimated the total costs of the ERP lifespan and distributed the costs according to the percentage cost structure identified in RQ1.

Both approaches were calculated by stepwise multiple regression. This method computes a prediction formula by adding and/or deleting candidate predictors until one 'best' subset of predictor variables is achieved. A formula was generated for each cost field.

These formulas should be evaluated by regarding their divergence (percentage error) from the average values of the actual sample.

The evaluation showed that cost fields are hard to predict, and the first approach does not deliver satisfying results. The percentage error, showing the average deviation between the estimated value and the actual value, is very high. This results in the disappointing realisation that the developed formula is not suited for accurately predicting costs.

Nevertheless, the errors appear to neutralise each other, which yields a good result for predicting the total costs of the ERP lifecycle. The second approach generated an MMRE of +/- 32%. Since the best identified value for ERP effort estimation was Widmer's 40%, this value seems to be an improvement on this research issue.

The developed formula for predicting the total costs is as follows:

Total Costs = 63,905 * x2 + 9,388 * x10 + 82,041 * x12 + 161,572 * x17 – 116,052 * x35

x2: No. of org. units or depts.
x10: No. of total users
x12: No. of modules[22]
x17: No. of external consultants
x35: Commitment management[23]

These five cost drivers seem to best predict the total costs. The four variables "x2: no. of org. units or depts.", "x10: no. of total users", "x12: no. of modules", "x17: no. of external consultants" and "x35: Commitment management" were expected to have a strong influence.
The strong impact of the variable "x35: commitment management" is an impressive result. The motivation and dedication of employees, being one of the five variables in the model, seem to be very important for the financial success of ERP projects.

In order to test the quality of the developed formula, the regression model was used to assess the suitability of the regressors and their premises. The examination yielded good results, which means that the formula is well suited to explain the total lifecycle costs.

---

[22] From: Purchasing, Sales, Material Management, CRM, SCM, Finance, Project Management, DMS, HR, Calculation, Production Planning, Detailed Production Planning, Capture of Production Data, Machine Data Collection, Personnel Accounting.

[23] Assessment between 1 to 10

7. CONCLUSION

The objective of this study was to make a contribution towards a better understanding of arising ERP costs. This thesis approached this research issue by combining the different costs types, the lifecycle phases and the cost drivers into one conceptual framework.
The findings yielded by this approach are summarised and discussed in this chapter. This chapter first provides an overview of the main findings. Then it shows its implications and relates to previous research.
Afterwards it reports the problems arising during the research. Finally, it shows its limitations and gives recommendations and perspectives for further research.

7.1 MAIN FINDINGS OF THE STUDY

This thesis yielded three main findings.
RQ1 found an average cost distribution which is presented in section 7.1.1. RQ2 identified the relevant cost drivers for each cost field. In total, it found 23 cost drivers which actually impact the ERP costs. Section 7.1.2 summarises these findings.
Swction 7.1.3 reports the findings of RQ3 which aimed to develop a formula for predicting ERP costs.
The findings are based on 72 eligible respondents which equates to a response rate of 3.8%.

7.1.1 FINDINGS RQ1

RQ1 explored the question “What are the costs of ERP Systems during their lifecycle phases“. The thesis discovered the average costs in absolute terms and in average percentage for each cost field during each lifecycle phase. The results are presented in the table below.

Table 7.1: Main finding I: Average cost structure of an ERP lifecycle

Lifecycle Phase	Cost Fields	Average Costs in absolute Terms	Average Percentage and Average Percentage Deviation	Average Percentage Deviation of each Lifecycle Phase
Evaluation Phase	y1 Internal Personnel Costs	35,855 EUR	3.7% ± 3.0%	5.8% ± 8.0%
	y2 External Personnel Costs	30,307 EUR	2.1% ± 2.6%	
Implementation Phase	y3 Internal Personnel Costs	136,637 EUR	9.8% ± 6.6%	41.1% ± 12.0%
	y4 External Personnel Costs	112,320 EUR	7.0% ± 5.0%	
	y5 ERP Software Costs	136,953 EUR	12.5% ± 10.1%	
	y6 Licence Costs	99,454 EUR	7.0% ± 5.1%	
	y7 Hardware Costs	58,250 EUR	4.8% ± 3.7%	
Maintanance Phase	y8 Internal Personnel Costs	390,280 EUR	14.8% ± 15.1%	53.2% ± 9.0%
	y9 External Personnel Costs	190,620 EUR	10.9% ± 9.5%	
	y10 ERP software costs	83,140 EUR	5.3% ± 5.9%	
	y11 Licence costs	225,550 EUR	16.3% ± 11.5%	
	y12 Hardware costs	95,830 EUR	5.9% ± 4.8%	

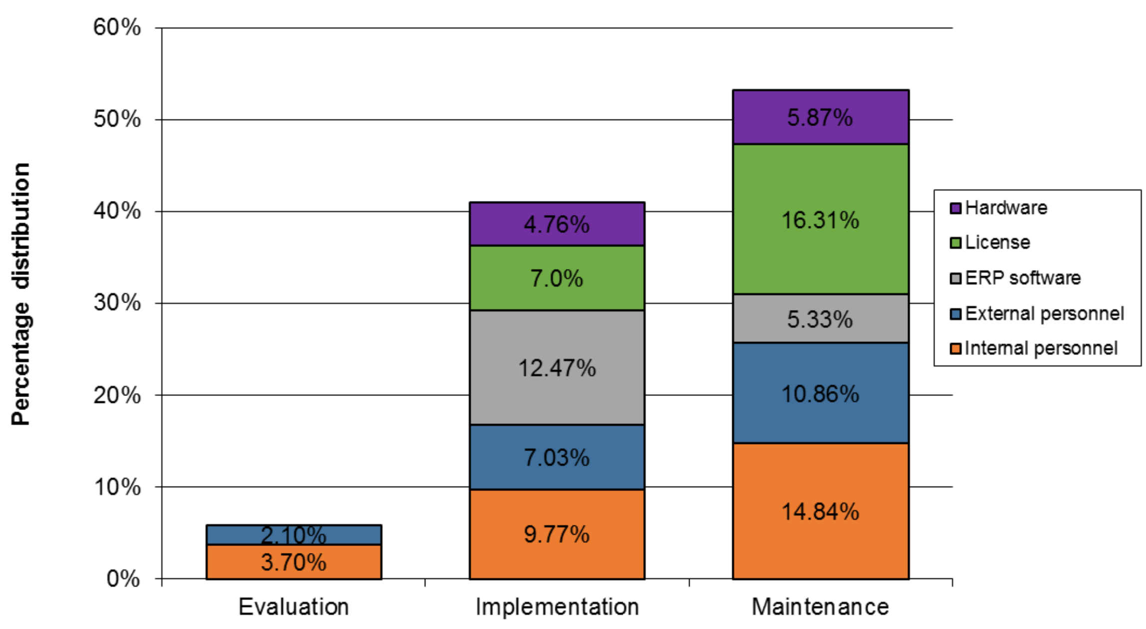
Table 7.1 shows that the maintenance phase regarded over 10 years is on average the one with the most arising costs. It requires 53.31% on average and ranges between 44.2 and 62.2% of the whole ERP lifecycle costs.
During this phase, the licence cost is the most dominant factor with 16.3% +/-11.5. This value is followed by internal personnel (14.8% +/- 15.1%) and external personnel (10.9% +/- 9.5%). The lowest expenses arise for hardware (5.9% +/- 4.8%) and software (5.33% +/- 5.9%) at this stage.

In total, the implementation phase costs on average 40.9% +/- 12% of the whole ERP lifecycle costs. By contrast, the software costs are the cost field with the highest costs during the implementation phase. They require 12.5 +/- 10.1% on average.
This expense is followed by internal personnel (9.8% +/- 6.6%). Both external personnel costs and licence costs are on average 7.0%. The average deviation for external personnel is +/- 5.0% and for licence +/- 5.1%. The lowest costs arise for hardware (4.8% +/- 3.7%) during the evaluation phase.

The lowest budget is needed for the evaluation phase. It requires between 0% and 16% of the total lifecycle costs. Internal personnel costs, with 3.7% +/- 3.0%, on average require a little more than the external personnel costs (2.1% +/-2.6%) of the total budget.

The figure below presents the average percentage cost structure graphically and underlines the maintenance phase as the phase with the highest expenses.

Figure 7.1: Average percentage cost distribution of the whole ERP lifecycle



These findings enable a very detailed insight of the average percentage cost distribution during the whole ERP lifecycle. The identifications create transparency about the costs arising over a time period of more than 10 years (evaluation + implementation + 10 years maintenance) and provide further knowledge about ERP costs.

7.1.2 FINDINGS RQ2

RQ2 aimed to answer the question “What are the cost drivers influencing ERP costs” and examined the responsible cost drivers for each cost field by correlation analysis.
The results are presented in the summary table below, which shows the correlating cost drivers and their correlation strength for each cost field.

Table 7.2: Main finding II: Correlation of cost fields and cost drivers

Cost field	Correlating Cost Drivers	Strength of correlation (r)
y1	x10 No. of total users	+0.408
Internal Personnel Costs	x1 No. of locations	+0.344
Evaluation Phase	x36 Satisfaction with ERP system	-0.309
	x31 Maturity of processes	-0.358
	x35 Commitment management	-0.390
	x33 Stability of organisation	-0.422
y2	x1 No. of locations	+0.406
External Personnel Costs	x10 No. of total users	+0.399
Evaluation Phase	x21 Team maturity	-0.346
	x31 Maturity of processes	-0.361
	x36 Satisfaction with ERP system	-0.426
	x33 Stability of organisation	-0.455
	x35 Commitment management	-0,460
y3	x10 No. of total users	+0.693
Internal Personnel Costs	x17 No. of external consultants	+0.643
Implementation Phase	x16 No. of internal project members	+0.526
	x2 No. of organisational units or depts.	+0.500
	x5-1 ERP system – SAP	+0.465
	x4 Revenue	+0.391
	x1 No. of locations	+0.368
	x5-6 ERP system – other	-0.334
	x33 Stability of organisation	-0.345
y4	x17 No. of external consultants	+0.719
External Personnel Costs	x10 No. of total users	+0.612
Implementation Phase	x16 No. of internal project members	+0.567
	x5-1 ERP system – SAP	+0.482
	x2 No. of organisational units or depts.	+0.443
	x12-9 Accounting module	+0.383
	x12-10 HRM module	+0.374
	x4 Revenue	+0.365
	x1 No. of locations	+0.356
	x33 Stability of organisation	-0.307
	x5-6 ERP system – other	-0.438
y5	x4 Revenue	+0.543
ERP Software Costs	x10 No. of total users	+0536
Implementation Phase	x17 No. of external consultants	+0.383
	x2 No. of organisational units or depts.	+0.346
	x1 No. of locations	+0.342
	x11 No. of user groups	+0.322

	x16 No. of internal project members	+0.322
	x33 Stability of organisation	-0.389
y6	x17 No. of external consultants	+0.519
Licence Costs	x1 No. of locations	+0.479
Implementation Phase	x16 No. of internal project members	+0.477
	x4 Revenue	+0.464
	x10 No. of total users	+0.440
	x2 No. of organisational units or depts.	+0.314
	x12-8 Finance module	+0.313
	x31 Maturity of processes	-0.347
y7	x2 No. of organisational units or depts.	+0.751
Hardware Costs	x5-1 ERP system – SAP	+0.441
Implementation Phase	x16 No. of internal project members	+0.431
	x4 Revenue	+0.426
	x10 No. of total users	+0.415
	x17 No. of external consultants	+0.385
y8	x10 No. of total users	+0.798
Internal Personnel Costs	x2 No. of organisational units or depts.	+0.627
Maintenance Phase	x17 No. of external consultants	+0,573
	x1 No. of locations	+0.533
	x4 Revenue	+0.472
	x5-1 ERP system – SAP	+0.468
	x9 No. of EDIs	+0.332
	x16 No. of internal project members	+0.325
	x6 No. of interfaces	+0.324
	x12-12 SCM module	+0.324
	x8 No. of reports	+0.323
	x5-6 ERP system – other	-0.315
	x33 Stability of organisation	-0.361
y9	x10 No. of total users	+0.778
External Personnel Costs	x17 No. of external consultants	+0.736
Maintenance Phase	x1 No. of locations	+0.560
	x16 No. of internal project members	+0.449
	x4 Revenue	+0.397
	x2 No. of organisational units or depts.	+0.392
	x5-1 ERP system – SAP	+0.391
	x12 No. of modules	+0.371
	x12-5 DMS module	+0.323
	x35 Commitment management	-0.301
	x33 Stability of organisation	-0.333
	x5-6 ERP system – other	-0.404

y10	x17 No. of external consultants	+0.442
ERP Software Costs	x16 No. of internal project members	+0.368
Maintenance Phase	x5-1 ERP system – SAP	+0.314
y11	x10 No. of total users	+0.655
Licence Costs	x17 No. of external consultants	+0.548
Maintenance Phase	x4 Revenue	+0.522
	x1 No. of locations	+0.466
	x2 No. of organisational units or depts.	+0.365
	x16 No. of internal project members	+0.359
	x9 No. of EDIs	+0.353
	x12-5 DMS module	+0.336
	x12 No. of modules	+0.324
	x11 No. of user groups	+0.317
	x12-9 Accounting module	+0.306
y12	x17 No. of external consultants	+0.569
Hardware Costs	x16 No. of internal project members	+0.428
Maintenance Phase	x10 No. of total users	+0.405
	x4 Revenue	+0.312

In total, this thesis identified 23 significant cost drivers. This results in the ability to reject some of the discussed ones. The systematic literature review discovered 35 potential cost drivers; this study is able to eliminate 20 of them as not significant.[24]

Besides identifying the relevant cost drivers influencing each cost field, the results provide a first starting point for developing a cost estimation model. The findings might enable the prediction of ERP costs. This approach will be specified in more detail in the next chapter.

7.1.3 FINDINGS RQ3

The purpose of RQ3 was to develop a model for predicting ERP costs. In accordance with the previous findings, this thesis pursued this aim through two different approaches.

[24] The two variables “x5: ERP system” and “x12: type of modules” have dummy variables: x5 has 6 and x12 has 15. The quantity of variables in total is 54.

The first approach was to estimate the costs of every single cost field. The second approach was to predict the total costs of the ERP lifespan and to distribute them to each cost field according to the cost structure identified in RQ1.

Both approaches were calculated by stepwise multiple regression. This method computes a prediction model by adding and/or deleting candidate predictors until one 'best' subset of predictor variables is achieved. A prediction model was generated for each cost field.

The results of RQ3 show that estimating the total costs of the ERP lifespan seems to be more precise and much easier than predicting the costs for each cost field. Neither the first nor the second approach delivered satisfying results for an accurate prediction of cost fields. The MMRE between the estimated values and the actual values are very high. However, the deviations seemed to compensate each other and thus yield a good result for predicting the total costs of the ERP lifecycle.
With an MMRE of +/- 32%, the second approach generated a better quality result than the first one.
In comparison to the best identified value for ERP effort estimation in the literature review, which was Widmer's MMRE of 40%, this value seems to contribute to the improvement in this research issue.

The formula developed for predicting the total costs is as follows:
Total Costs = 63,905 * x2 + 9,388 * x10 + 82,041 * x12 + 161,572 * x17 – 116,052 * x35

The predicting variables are as follows:
x2: No. of org. units or depts.
x10: No. of total users
x12: No. of modules[25]
x17: No. of external consultants
x35: Commitment management[26]

---

[25] From: Purchasing, Sales, Material Management, CRM, SCM, Finance, Project Management, DMS, HR, Calculation, Production Planning, Detailed Production Planning, Capture of Production Data, Machine Data Collection, Personnel Accounting.

With an adjusted R^2 variance of 92.2, these independent variables are able to explain 92.8% of the result. Hence, this formula seems to be a good tool to predict the costs of an ERP lifespan.

7.2 IMPLICATIONS TO PRACTICE

The literature review identified a lack of suitable models to predict the costs arising in ERP projects. Consequently, in practice, many organisations need to rely on the quotations of different vendors since they cannot verify the calculated costs stated in the offers. The introduction reported cost overruns of 178%, so apparently there is a mismatch between the vendor quotation and the final costs incurred.

One reason for that could be that vendor invoicing is in accordance with the actually accomplished expenditure. Since suppliers compete against each other in order to secure an order, they may calculate their quotations optimistically. So far, it has not been possible to check this calculation.

The results of this thesis might be used to equip ERP managers in organisations with concrete tools to predict their costs.

Firstly, the identified average cost structure of ERP projects provides transparency of the emerging costs. It gathers all possible cost types incurred and identifies their average expenses in percent. This might lead to more precise calculations just because one is aware of them and does not forget an important part.

Internal personnel costs might be such a factor bound to be underestimated. During the evaluation phase, the internal personnel costs are higher than the expenditures for external personnel. Requiring the second highest expenses during the implementation and evaluation phase, the internal personnel costs are an important part of the emerging costs which should not be neglected.

Secondly, the cost structure gives an overview of which costs are expected to arise at what stage of the lifecycle. This might increase the awareness of ERP managers and in-

[26] Assessment between 1 to 10

fluence their budget planning. The structure clearly highlights the maintenance costs as the biggest cost part during the ERP lifecycle. Requiring 53% +/- 9.0% on average, the maintenance costs are on average more than 10% higher than the costs arising during the implementation phase. Since previous research focused more on the implementation, the costs during the maintenance phase might be a disregarded factor in many predictions.
However, it does in fact require the highest costs of the whole lifespan, which is an important fact for the provision of capital.

Thirdly, the cost structure might help ERP managers to check quotations of vendors for their plausibility. At the very beginning of an ERP project, after evaluating the basic needs, vendors submit a tender. These are mainly based on three positions: licence costs (implementation), maintenance cost for licence, and costs for their employees, like consultants, programmers, etc. (external personnel costs).
In many ERP projects, licence costs are the most predictable and reliable variable. They depend only on the number of users, in most cases. The only ways to increase the tender is either to enlarge the number of users or raise the prices. Therefore, this variable seems to be an accurate factor for derivation.
Knowing that licence costs during the implementation phase require on average 7.0% of the total ERP costs, ERP managers could construe the costs of each cost field and the total ERP costs. The same procedure can be applied to licence costs for maintenance.
Furthermore, managers can check the relation of licence costs and external personnel costs. This could give a first indication whether the projected amount of external personnel costs provides a realistic picture.

Fourthly, the identification of the relevant cost drivers of each cost field in RQ2 enables ERP managers to focus on them. They might be sensitive to the factors which can easily increase the costs at all stages of the lifecycle in order to keep the costs within the budget. This awareness might improve the planning and controlling performances.

Fifthly, the developed formula in RQ3 provides a tool for ERP managers to predict the costs of the ERP project up-front. The formula is rather easy to calculate, which enables all organisations to apply it. That means that they do not have to rely on the vendors' quotations any longer but can check the required budget for the lifecycle themselves.

Of course, the formula deviates by +/- 32% on average, but considering cost overruns by an average of 178%, this seems a moderate deviation which at least guides the budget scope.

Nevertheless, ERP projects are very vast and complex undertakings. Since the different requirements and structures of ERP projects in different organisations differ extremely, an accurate prediction of costs seems to be hardly possible.
The deviation of the prediction formula seems to support this argument.

7.3 RELATIONSHIP TO PREVIOUS RESEARCH

The systematic literature review identified a very fragmented research field which seems to be rather unexplored. Aside from the fact that no suitable model for ERP cost estimation exists, three main weaknesses in ERP cost estimation approaches were identified by the review.

The definition of costs in almost all these approaches is a quite opaque issue. There is a fragmented understanding of costs incurred. Since most researchers do not state their meaning of costs in relevant ERP effort estimation papers, the remaining papers do not present a consensual understanding. Every researcher focuses on a different aspect and no one has ever mapped an overall picture of all incurred ERP cost before. All found literature presents only extracts.

The second weakness is the fragmented regard of potential cost driver candidates. Every research makes other suggestions about cost-causing factors. The literature shows a picture quite similar to that presented for cost types: research does not provide an overall picture of these influencing factors. Almost every researcher focuses on single factors.

Another important aspect of cost estimation is the time at which costs occur. Surprisingly, this aspect has not been given much attention in the previously developed approaches. ERP systems typically go through different lifecycle phases during their lifetime. The literature review found that all research mainly considers only the implementation phase. The costs incurred during other phases are completely ignored in most studies. Since the consideration of maintenance and evaluation is an important factor for the total costs, disregarding this step could be fatal for ERP prediction concepts.

This thesis is based on the previous research and builds a research concept grounded on the identified findings. The research concept synthesises the three above-mentioned aspects “cost types”, “cost drivers” and “lifecycle phases” into one framework and pools all the identified suggestions.
First, it identified the five different cost types “internal personnel”, “external personnel”, “hardware”, “software” and “licence”. These different cost types have not been synthesised into one model before. Then, these cost types are matched to the three identified lifecycle phases “evaluation”, “implementation” and “maintenance”, which results in twelve different cost fields.

These cost fields should be examined for their influencing factors, the so-called cost drivers. To this end, this thesis pooled all suggested cost drivers from previous researches. Many cost drivers are in debate within the literature. Some have been validated in empirical tests; others originated from theoretical frameworks or different cost estimation models and have not been empirically confirmed before. A lot of assumptions and speculation appear within the cost estimation of ERP systems.
In total, the systematic literature review identified 64 cost driver candidates in previous researches. After selection, 35 were considered to be relevant in this study. All of them were analysed for their correlation with each cost field. 23 cost drivers could be verified in this study and 20 of them could be rejected.[27] Subsequently, these findings were used to develop the prediction formula.
The pooling of all relevant cost types allows for more transparency of possible emerging costs and might increase the awareness of them. This differentiated consideration of ERP costs enables a division of the total costs into smaller entities. This allows a better planning and control of ERP costs and permits a better understanding of costs incurred in an ERP project.

[27] The two variables “x5: ERP system” and “x12: type of modules” have dummy variables: x5 has 6 and x12 has 15. The total number of variables is 54.

Altogether, the developed conceptual framework builds up several contributions to theoretical knowledge in regard to the topic of ERP effort estimation. Through the combination of the three dimensions “lifecycle phases”, “cost types”, and “cost drivers”, the conceptual framework describes the cost structure and the factors influencing it during the whole ERP lifecycle. While previous research concentrated only on extracts of ERP costs, this conceptual framework considers the whole ERP lifecycle (and not only its implementation) as well as several different cost aspects (and not only external personnel costs, for example).

The percentual cost structure of an ERP lifecycle according to the 12 cost fields is therefore the first contribution to theoretical knowledge. The relevance to practice has already been discussed in section 7.2.

The second contribution to knowledge is the validation and allocation of potential cost drivers to the single cost fields. While previous research made no distinction between ERP costs and when they occur during the ERP lifecycle, this research established a link between the cost drivers and 12 cost fields. This insight can also help researchers developing ERP cost estimation models from other disciplines than correlation and regression analysis. The literature review has shown that, in all models, the most important issue is the right input variables.

The third contribution to knowledge is the development of the estimation formula. With an MMRE of +/- 32%, this estimation formula achieves the currently best results in ERP effort estimation.

7.4 LIMITATIONS

This chapter considers the criticism anticipated with regard to this thesis and reports on the limitations identified the research design and the analysis of the results. Each topic is described individually in the sections below.

Research Design

Initially, this thesis was crafted in accordance with the postpositivist paradigm. This paradigm guided this research in ontology, epistemology and methodology. Other paradigms and their approaches are not considered and might have led to other results.

The consequently chosen quantitative research design did not support the finding of undiscovered cost types and cost drivers, as already mentioned in the section “framework” above. The closed-question design did not allow respondents to include new points or ideas. It is possible that important additional information did not have any influence on the results of this thesis. Moreover, the design of the questionnaire disregards possible specific features of organisations.

An additional limitation of the quantitative design is that the transformation of variables in Likert scales might not measure what the researcher had intended with regard to the inquired. To attenuate that risk, the questionnaire was discussed with five companies in advance, but there a certain risk of inaccuracy remains.

Furthermore, the target group and sample technique involve some restrictions. The findings of the study are restricted to German SMEs with 30 – 1,500 employees having their core business the industrial sector. It is questionable whether the results of this thesis can be transferred to other business sectors, other countries and companies with different employee cultures and structures.

The main weakness of this study is the sample technique. Since a self-selecting sample was applied, the respondents might not be representative. To ensure this, a random sampling method should have been applied. Unfortunately, this was not possible within the research. Consequently, this research has its limitations in terms of external validity and generalisability.
Furthermore, the predictive validity could not be accessed in this study. Reconducting the survey a few years later would have exceeded the scope of this thesis. This could be a topic for further research.

Analysis of Results

This thesis has a further limitation in the assessment of the model quality of ERP effort estimation.

In RQ2, this thesis addresses only relationships between variables, namely between each single cost field and the cost drivers. Some of these relationships were anticipated while others were astonishing, and others again were expected to have a much stronger influence.

Although this thesis provides some approaches to an explanation, the relationships were not tested for their causality. Consequently, the question why the identified cost drivers have an influence remains unsolved. The causality analysis would have exceeded the scope of this thesis, but it might be a topic for further studies.

In RQ3, the quality of results is evaluated by regarding the quality criterion MMRE. This criterion was selected because it is the only one stated in all of the identified empirical studies. Myrtveit & Stensrud (1999) did not consider pred (0.2) as a quality criterion. Widmer (2004) did not state the MdMRE. To be able to compare the findings, this thesis focuses on the MMRE. Different quality criteria might have led to different outcomes and interpretations.
However, the MMRE of the different cost fields compensate each other and create a good result for the total costs of an ERP project. This might be an indication for shifts between the cost fields and lifecycle phases, which indicates a possible dependence between the single cost fields as well as between the different phases of the lifecycle.
This thesis gives explanation approaches for that, but does not examine these deferrals or dependences. This might be a starting point for further research.

7.5 RECOMMENDATIONS

The rationale of this study was that the creation of small cost entities, namely the 12 cost fields, increases transparency and allows for more accurate cost predictions.
The developed prediction formulas showed that estimating the total costs of the ERP lifespan is more accurate than predicting the costs for each cost field. Estimating the costs for the different cost fields yields very high MMRE values between the estimated and the actual costs. Nevertheless, the deviations seem to compensate each other and produce a good result for predicting the total costs of the ERP lifecycle.
Apparently, strong shifts between the single cost fields took place. The analysis of these shifts could be a possible area for future research in order to improve the cost field prediction. Qualitative research design might be an approach to access this field. Future studies could examine why shifts occur and suggest how they could be integrated into future prediction formulas for estimating the costs of each cost field.

Another possible area for future research is the optimisation of the prediction formula. Even if RQ3 provided a calculation tool which allows a forecast accuracy of +/- 32% which is bound to provide improvement, this result is not perfect. Reducing the estimation deviation in order to improve the reliability of results seems a worthwhile research topic to explore.
Using the findings for other model developments could also be a research topic for optimising the cost estimation. These results were generated by applying regression analysis. One approach could be to use the sample and the verified cost drivers in analogy-based designs and to compare these findings to the results of this research. It might be possible that other methods yield more precise results.

A future research topic could be the causality testing of the relationships identified between the cost fields and the cost drivers. As mentioned in the previous chapter, this thesis addresses only the relationship between cost fields and the cost drivers and did not test their causality. The causality analysis might be a topic for further studies.

Research on the transferability and generalisability of the findings to other organisations, other business sectors and other countries could be another avenue for future studies. Since the findings of this study were based on German SMEs with 30 – 1,500 employees having their core business the industrial sector, it is questionable whether the results of this thesis can be transferred to other business sectors, other countries and companies with a higher number of employees and more complex structures. Other cultures might have a totally different approach to ERP projects. It could be very interesting to compare those findings to the findings of this thesis.

The last recommendation for further research is the testing of the predictive validity. As mentioned in the previous chapter, this could not be done within the scope of this study. Reconducting the survey in a few years in order to examine the equality of results could be another topic for further research.

8. BIBLIOGRAPHY

Aamodt, A., & Plaza, E. (1994). Case-based reasoning: Foundational issues, methodological variations, and system approaches. *AI communications, 7*(1), 39-59.

Abts, C., Boehm, B. W., & Clark, E. B. (2000). *COCOTS: A COTS software integration lifecycle cost model-model overview and preliminary data collection findings.* Paper presented at the Proceedings of ESCOM-SCOPE Conference.

Abts, D., & Mülder, W. (2009). *Grundkurs Wirtschaftsinformatik: eine kompakte und praxisorientierte Einführung*: Springer-Verlag.

Al-Mashari, M. (2003). Enterprise resource planning (ERP) systems: a research agenda. *Industrial Management & Data Systems, 103*(1), 22-27. doi: 10.1108/02635570310456869

Al-Mashari, M., Al-Mudimigh, A., & Zairi, M. (2003). Enterprise resource planning: a taxonomy of critical factors. *European journal of operational research, 146*(2), 352-364.

Arb, R. v. (1997). *Vorgehensweisen und Erfahrungen bei der Einführung von Enterprise-Management-Systemen dargestellt am Beispiel von SAP R/3.* University Bern, Bern. Retrieved from online internet research. URL:<http://www.digital-publications. ch/vonarb/vonarb. zip> [12.02. 2006]

Backhaus, K., Erichson, B., Plinke, W., & Weiber, R. (2008). *Multivariate Analysemethoden: Eine anwendungsorientierte Einführung* (12 ed.). Berlin: Springer.

Baddeley, M., & Barrowclough, D. (2009). *Running regressions*: Cambridge University Press.

Barki, H., & Pinsonneault, A. (2002). Explaining ERP implementation effort and benefits with organizational integration.

Barki, H., & Pinsonneault, A. (2003). The construct of organizational integration: A research framework and its application to enterprise systems research. *Cahier du GReSI no, 3*, 04.

Barki, H., & Pinsonneault, A. (2005). A model of organizational integration, implementation effort, and performance. *Organization Science, 16*(2), 165-179.

Baruch, Y., & Holtom, B. C. (2008). Survey response rate levels and trends in organizational research. *Human Relations, 61*(8), 1139-1160.

Bell, E., & Bryman, A. (2007). The ethics of management research: an exploratory content analysis. *British Journal of Management, 18*(1), 63-77.

Bell, J. (2005). *Doing your research project* (4 ed.). Buckingham: Open University Press.

Bingi, P., Sharma, M. K., & Godla, J. K. (1999). Critical Issues Affecting an ERP Implementation. *Information Systems Management, 16*(3), 7-14. doi: 10.1201/1078/43197.16.3.19990601/31310.2

Bititci, U. S., & Carrie, A. S. (1998). *Strategic management of the manufacturing value chain* (Vol. 2): Springer.

Blumberg, B., Cooper, D., & Schindler, P. (2005). Survey Research. *Business research methods*, 243-276.

Brehm, L., Heinzl, A., & Markus, M. (2001, 3-6 January). *Tailoring ERP Systems: A Spectrum of Choices and their Implications.* Paper presented at the PROCEEDINGS OF THE ANNUAL HAWAII INTERNATIONAL CONFERENCE ON SYSTEM SCIENCES.

Buckhout, S., Frey, E., & Nemec, J. (1999). Making ERP succeed: turning fear into promise. *Strategy and Business*, 60-73.

Burns, R. P., & Burns, R. (2008). *Business research methods and statistics using SPSS*: Sage.

Camanho, A. (2007). Performance Measurement using Data Envelopment Analysis (DEA) and Malmquist indices: Issues, Challenges and Applications: University of Porto, Portugal.

Chang, S.-I., Gable, G., Smythe, E., & Timbrell, G. (2000). *A Delphi examination of public sector ERP implementation issues.* Paper presented at the Proceedings of the twenty first international conference on Information systems.

Chung, S., & Synder, C. (1999). *ERP initiation-a historical perspective*. Paper presented at the AMCIS 1999 Proceedings.

Cohen, J., Cohen, P., West, S. G., & Aiken, L. S. (2013). *Applied multiple regression/correlation analysis for the behavioral sciences*: Routledge.

Creswell, J. W. (2009). *Research design: Qualitative, quantitative, and mixed methods approaches*: Sage.

Daneva, M. (2007). *Approaching the ERP project cost estimation problem: An experiment.* Paper presented at the Proceedings of the First.

Daneva, M. (2008a). *Complementing approaches in ERP effort estimation practice: an industrial study.* Paper presented at the Proceedings of the 4th international workshop on Predictor models in software engineering.

Daneva, M. (2008b). Integrating Portfolio Management and Simulation Concepts in the ERP Project Estimation Practice *Requirements Engineering: Foundation for Software Quality* (pp. 147-152): Springer.

Daneva, M. (2008c). Managing Uncertainty in ERP Project Estimation Practice: An Industrial Case Study *Product-Focused Software Process Improvement* (pp. 90-104): Springer.

Daneva, M., Wettflower, S., & de Boer, S. (2008). Uncertainty in ERP Effort Estimation: A Challenge or an Asset? *Software Process and Product Measurement* (pp. 208-222): Springer.

Daneva, M., & Wieringa, R. (2008). Cost Estimation for Cross-Organizational ERP Projects: Research Perspectives. *Software Quality Journal, 16*, 459-481.

Davenport, T. H. (1998). Putting the enterprise into the enterprise system. *Harvard Business Review., 76*(4), 121-131.

Davenport, T. H. (2000). *Mission critical : realizing the promise of enterprise systems*. Boston, MA: Harvard Business School Press.

Davis, A. (2005). *ERP Customization Impacts on the Strategic Alignment and System Quality.* Paper presented at the Southern Association of Information Systems Conference.

Diekmann, A. (2004). *Empirische sozialforschung. Grundlagen, Methoden, Anwendungen* (12 ed.): Rowohlt Taschenbuch Verlag.

Dillman, D. A. (2000). *Mail and internet surveys: The tailored design method* (Vol. 2): Wiley New York.

Easterby-Smith, M., Thorpe, R., & Jackson, P. (2008). *Management research* (3 ed.): Sage Publications.

Eck, D., Brundick, B., Fettig, T., Dechoretz, J., & Ugljesa, J. (2009). Parametric estimating handbook. *The International Society of Parametric Analysis (ISPA)*.

Endriss, U. (2011). Logic and social choice theory. *Logic and philosophy today, 2*, 333-377.

Equey, C., Kusters, R. J., Varone, S., & Montandon, N. (2008). *Empirical Study of ERP Systems Implementation Costs in Swiss SMEs.* Paper presented at the ICEIS (1).

Erkan, T. E. (2009). BPR Effect on ERP Implementation: a Comparative Case Study. *World Academy of Science, Engineering and Technology, 30*, 1508-1511.

Esteves, J., Pastor-Collado, J., & Casanovas, J. (2002). *MONITORING BUSINESS PROCESS REDESIGN IN ERP IMPLEMENTATION PROJECTS*. Paper presented at the AMCIS 2002 Proceedings.

Esteves, J., & Pastor, J. (1999). *An ERP lifecycle-based research agenda.* Paper presented at the Published in: First International workshop in Enterprise Management and Resource. Planning: Methods, Tools and Architectures–EMRPS.

Fishburn, P. C. (1977). Condorcet social choice functions. *SIAM Journal on Applied Mathematics, 33*(3), 469-489.

Francalanci, C. (2001). Predicting the implementation effort of ERP projects: empirical evidence on SAP/R3. *Journal of information technology, 16*(1), 33-48. doi: 10.1080/02683960122849

Glass, R. L. (1998). Enterprise Resource Planning - Breakthrough and/or Term Problem? *ACM SIGMIS Database, 29*(2), 13-16.

Gliner, J. A., Morgan, G. A., & Leech, N. L. (2009). *Research methods in applied settings: An integrated approach to design and analysis* (2 ed.). New York: Taylor & Francis US.

Grabot, B., Mayère, A., & Bazet, I. (2008). *ERP systems and organisational change a socio-technical insight*. London: Springer.

Grant, D. (2002). A wider view of business process reengineering. *Communications of the ACM, 45*(2), 85-90.

Gronau, N. (2010). *Enterprise Resource Planning: Architektur, Funktionen und Management von ERP-Systemen* (2. ed.). Frankfurt/Oder: Oldenbourg.

Guba, E. G. (1990). *The paradigm dialog*: Sage.

Hammer, M., & Champy, J. (1993). *Reengineering the corporation : a manifesto for business revolution*. New York, NY: HarperBusiness.

Hansen, T. (2006). *Multidimensional effort prediction for ERP system implementation.* Paper presented at the On the Move to Meaningful Internet Systems 2006: OTM 2006 Workshops.

Harwood, S. (2003). *ERP: The implementation cycle*: Butterworth-Heinemann.

Hesseler, M., & Görtz, M. (2007). *Basiswissen ERP-Systeme. Auswahl, Einführung & Einsatz betriebswirtschaftlicher Standardsoftware* (1 ed.). Herdecke: W3L.

Holland, C., Light, B., & Kawalek, P. (1999). *Beyond enterprise resource planning projects: innovative strategies for competitive advantage.* Paper presented at the Proceedings of the 7th International Conference on Information Systems, Copenhagen, Denmark.

Israel, G. D. (1992). *Determining sample size*: University of Florida Cooperative Extension Service, Institute of Food and Agriculture Sciences, EDIS.

Janssens, G., Kusters, R., & Heemstra, F. (2008). Sizing ERP implementation projects: an activity-based approach. *International Journal of Enterprise Information Systems (IJEIS), 4*(3), 25-47.

Kalton, G. (1983). *Introduction to survey sampling* (Vol. 7): SAGE Publications, Incorporated.

Keung, J. (2009). *Software development cost estimation using analogy: a review.* Paper presented at the Software Engineering Conference, 2009. ASWEC'09. Australian.

Kitchenham, B. (2004). Procedures for performing systematic reviews. *Keele, UK, Keele University, 33*, 2004.

Klamler, C. (2005). On the closeness aspect of three voting rules: Borda–Copeland–Maximin. *Group Decision and Negotiation, 14*(3), 233-240.

Klaus, H., Rosemann, M., & Gable, G. G. (2000). What is ERP? *Information Systems Frontiers, 2*(2), 141-162.

Koch, S. (2007). ERP implementation effort estimation using data envelopment analysis *Technologies for business information systems* (pp. 121-132): Springer.

Koch, S., & Mitlöhner, J. (2010). Effort estimation for enterprise resource planning implementation projects using social choice–a comparative study. *Enterprise Information Systems, 4*(3), 265-281.

Kolodner, J. L. (1992). An introduction to case-based reasoning. *Artificial Intelligence Review, 6*(1), 3-34.

Kuhn, J. (2003). Workbook for Statistics 512 - Applied Regression Analysis Purdue University.

Kurbel, K. (2009). Enterprise Resource Planning and Integration *Encyclopedia of Information Science and Technology, Second Edition* (pp. 1398-1404): IGI Global.

Kurbel, K. (2011). *Enterprise Resource Planning und Supply Chain Management in der Industrie* (7 ed.). München: Oldenbourg Wissenschaftsverlag.

Kusters, R. J., Heemstra, F. J., & Jonker, A. (2007). *Determining the costs of ERP implementation.* Paper presented at the ICEIS (1).

Kwon, S. B., & Shin, K.-s. (2005). PPSS: CBR system for ERP project pre-planning *Artificial Intelligence and Simulation* (pp. 157-166): Springer.

Lienert, G. A., & Raatz, U. (1998). *Testaufbau und Testanalyse* (6 ed.). Weinheim: Psychologie Verlags Union.

Lincoln, Y. S. (1990). The making of a constructivist: A remembrance of transformations past. In E. Guba (Ed.), *The paradigm dialog* (pp. 258-276): SAGE Publications Inc.

Lincoln, Y. S., & Guba, E. (1985). *Naturalistic inquiry*. Beverly Hills, California: SAGE Publications Inc.

Markus, M. L., Axline, S., Petrie, D., & Tanis, S. C. (2000). Learning from adopters' experiences with ERP: problems encountered and success achieved. *Journal of information technology, 15*(4), 245-265.

Markus, M. L., & Tanis, C. (2000). The enterprise systems experience–from adoption to success. *Framing the domains of IT research: Glimpsing the future through the past, 173*, 207-173.

Mauterer, H. (2002). *Der Nutzen von ERP-Systemen*: Springer.

Mische, M. A. (2001). Defining Systems Integration. In J. Myerson (Ed.), *Enterprise Systems Integration* (pp. 3-10): Auerbach.

Moosbrugger, H., & Kelava, A. (2012). Qualitätsanforderungen an einen psychologischen Test (Testgütekriterien) *Testtheorie und Fragebogenkonstruktion* (pp. 7-26): Springer.

Mukhopadhyay, T., Vicinanza, S. S., & Prietula, M. J. (1992). Examining the feasibility of a case-based reasoning model for software effort estimation. *MiS Quarterly*, 155-171.

Myrtveit, I., & Stensrud, E. (1999a). *Benchmarking COTS projects using data envelopment analysis.* Paper presented at the Software Metrics Symposium, 1999. Proceedings. Sixth International.

Myrtveit, I., & Stensrud, E. (1999b). A controlled experiment to assess the benefits of estimating with analogy and regression models. *Software Engineering, IEEE Transactions on, 25*(4), 510-525.

Ng, J. K. C., Ip, W. H., & Lee, T. C. (1999). A paradigm for ERP and BPR integration. *International Journal of Production Research, 37*(9), 2093-2108. doi: 10.1080/002075499190923

Oudshoorn, I. (2004). Development of Packaged Software. *Vrije Universiteit Amsterdam.*

Parr, A. N., & Shanks, G. (2000, 4-7 Jan.). *A Taxonomy of ERP Implementation Approaches.* Paper presented at the Proceedings of the 33rd Annual Hawaii International Conference on System Sciences. .

Patnaik, K., Malhotra, S., & Sahoo, B. (2004). Software Development Effort Estimation using CBR: A Review.

Phillips, D. (1987). On what scientists know and how they know it. In W. R. Shadish & C. S. Reichardt (Eds.), *Evaluation Studies Review Annual* (Vol. 12, pp. 377-399): Sage Publications.

Radovilsky, Z. (2004). Enterprise Resource Planning (ERP). In H. Bidgoli (Ed.), *The Internet Encyclopedia* (Vol. 1, pp. 707-717): John Wiley & Sons.

Rammstedt, B. (2004). Zur Bestimmung der Güte von Multi-Item-Skalen: Eine Einführung. *ZUMA How-to-Reihe, 12.*

Risse, M. (2005). Why the count de Borda cannot beat the Marquis de Condorcet. *Social Choice and Welfare, 25*(1), 95-113.

Ross, J. W. (1999). *The ERP revolution: surviving vs. thriving*. INformation Systems Frontiers, (2, 2).

Saunders, M., Lewis, P., & Thornhill, A. (2007). *Research Methods For Business Students* (4 ed.). Essex: Pearson Education Limited.

Schniederjans, M. J., & Kim, G. C. (2003). Implementing enterprise resource planning systems with total quality control and business process reengineering: survey results. *International Journal of Operations & Production Management, 23*(4), 418-429.

Schwandt, T. R. (1990). Paths to inquiry in the social disciplines: Scientific, constructivist, and critical theory methodologies. In E. Guba (Ed.), *The paradigm dialogue* (pp. 258-276).

Sen, A. (1986). Social choice theory. *Handbook of mathematical economics, 3*, 1073-1181.

Sheehan, K. B. (2001). E-mail Survey Response Rates: A Review. *Journal of Computer-Mediated Communication, 6*(2), 0-0. doi: 10.1111/j.1083-6101.2001.tb00117.x

Shehab, E. M., Sharp, M. W., Supramaniam, L., & Spedding, T. A. (2004). Enterprise resource planning: An integrative review. *Business Process Management Journal, 10*(4), 359-386. doi: 10.1108/14637150410548056

Shepperd, M., & Schofield, C. (1997). Estimating software project effort using analogies. *Software Engineering, IEEE Transactions on, 23*(11), 736-743.

Shepperd, M., Schofield, C., & Kitchenham, B. (1996). *Effort estimation using analogy.* Paper presented at the Proceedings of the 18th international conference on Software engineering.

Sherman, H. D., & Zhu, J. (2006). *Service productivity management: Improving service performance using data envelopment analysis (DEA)*: Springer.

Somers, T. M., & Nelson, K. G. (2004). A taxonomy of players and activities across the ERP project life cycle. *Information & Management, 41*(3), 257-278.

Stefanou, C. (2001). A framework for the ex-ante evaluation of ERP software. *European Journal of Information Systems, 10*(4), 204-215.

Stensrud, E. (2001). Alternative approaches to effort prediction of ERP projects. *Information and software technology, 43*(7), 413-423.

TACSO. (2010). The Integrated Organisation Model. Retrieved from http://www.tacso.org/Capacity_Development/online_courses/cso_management_course/csomt_03.pdf

Tashakkorie, A., & Teddlie, C. (1998). *Handbook of Mixed Methods*. London: SAGE Publications Ltd.

Walkerden, F., & Jeffery, R. (1999). An empirical study of analogy-based software effort estimation. *Empirical software engineering, 4*(2), 135-158.

Welker, M., Werner, A., & Scholz, J. (2005). *Online-Research: Markt-und Sozialforschung mit dem Internet*: dpunkt-Verlag.

Widmer, T. (2004). *Schätzung und Beeinflussung der Kosten von ERP-Systemen in Schweizer KMU.* Swiss Federal Institute of Technology Zurich.

Yen, T. S., Idrus, R., & Yusof, U. K. (2011). A FRAMEWORK FOR CLASSIFYING MISFITS BETWEEN ENTERPRISE RESOURCE PLANNING (ERP) SYSTEMS AND BUSINESS STRATEGIES. *Asian Academy of Management Journal, 16*(2), 53-75.

Young, P. (1995). Optimal voting rules. *The Journal of Economic Perspectives, 9*(1), 51-64.

APPENDIX I: NUMERIC REFERENCES OF INCLUDED STUDIES

1. Arb, R. von (1997). Vorgehensweisen und Erfahrungen bei der Einführung von Enterprise-Management-Systemen, PhD. Dissertation Universität Bern, Switzerland

2. Equey, C., Kusters, R.J., Varone, S., & Motandon, N. (2008). Empirical study of ERP system costs in Swiss SMEs. *Proceedings of the Tenth International Conference on Enterprise Information Systems.* ICEIS 2008*,* 143-148

3. Francalanci, C. (2001). Predicting the Implementation Effort on ERP projects, Journal of Information Technology, 16(1), 33-48

4. Kusters, R.J., Heemstra, F.J., & Jonker, A. (2007). Determining the costs of ERP Implementation; Proceedings of the 9th International Conference on Enterprise Information Systems, Vol. Database and Information Systems Integration, 102-110

5. Barki, H., Pinsonneault, A. (2002). Explaining ERP Implementation Effort and Benefits with Organisational Integration. Computer and Information Science, Volume 01, Issue 02, 1 -27

6. Stensrud, E. (2001). Alternative Approaches to Effort Prediction of ERP Projects, Information and Software Technology, 43, 413-423

7. Widmer, T. (2004). Schätzung und Beeinflussung der Kosten von ERP-Systemen in Schweizer KMU (Translation: Estimation of and influences on ERP system costs in Swiss SMEs), Dissertation Eidgenössische Technische Hochschule Zürich, Switzerland

8. Koch, S., Mitlöhner, J. (2010). Effort estimation for enterprise resource planning implementation projects using social choice – a comparative study. International Journal of Enterprise Information Systems, 265 – 281

9. Daneva, M. (2010). Balancing uncertainty of context in ERP project estimation: an approach and a case study. Journal of Software Maintenance and Evolution: Research and Practice, 329 – 357

10. Janssen, G., Kusters, R., Heemstra, F. (2008). Sizing ERP implementation projects: an activity-based approach. International Journal of Enterprise Information Systems, 25 – 47

11. Daneva, M. (2008). Integrating Portfolio Management and Simulation Concepts in the ERP Project Estimation Practice. Requirements Engineering: Foundation for Software Quality, 147 – 152

12. Daneva, M. (2008). Managing Uncertainty in ERP Cost Estimation Practice. An Industrial Case Study. Product-focused software process improvement, 90 – 104

13. Daneva, M., Wieringa, R. (2008). Cost estimation for cross-organizational ERP projects: research perspectives. Software quality Journal, 459 – 481

14. Daneva, M. (2007). Approaching the ERP Project Coast Estimation Problem: An Experiment. International Symposium on Empirical Software Engineering and Measurement, IEEE Computer Society Press, 500

15. Daneva, M., Wettflower, S., de Boer, S. (2008). Uncertainty in ERP Effort Estimation: A Challenge or an Asset? Software Process and Product Measurement, 208 – 222

16. Hansen, T., (2006). Multidimensional effort prediction for ERP system implementation. ON THE MOVE TO MEANINGFUL INTERNET SYSTEMS 2006: OTM 2006 WORKSHOPS, PT 2, PROCEEDINGS, 1402 – 1408

17. Barki, H., Pinsonneault, A. (2005). A model of organizational integration, implementation effort, and performance. Organization Science 165 – 179

18. Kwon, S.B., Shin, K.S. (2004). PPSS: CBR system for ERP project pre-planning. Artificial intelligence and simulation, 157 – 166

19. Stensrud, E., Foss, T., Kitchenham, B., Myrtveit, I. (2003). A further empirical investigation of the relationship between MRE and project size. Empirical Software Engineering 139 – 161

20. Myrtveit, I., Stensrud, E. (1999). A controlled experiment to assess the benefits of estimating with analogy and regression models. IEEE Transactions on Software Engineering, 510 – 525

21. Myrtveit, I., Stensrud, E., Olsson, U. (2000). Assessing the benefits of imputing ERP projects with missing data. 7th International Software Metrics Symposium – Metrics 2001 Proceedings, 78 – 84

APPENDIX II: EXCLUSION OF STUDIES WITHIN SYSTEMATIC LITERATURE REVIEW

Author	Title	Year	Exclusion Criteria
Ford, A., Brown, D.	How SMEs are accessing and using enterprise software: the experience of UK-based SMEs from choice to evaluation.	2010	2, 3
Celar, S., Mudnic, E., Katajdzic, E.	Software size estimating method based on MK II FPA 1,3 unadjusted.	2009	2, 3
Osei-Bryson, K. M., Dong, L. Y., Ngwenyama, O.	Exploring managerial factors affecting ERP implementation: an investigation of the Klein-Sorra model using regression splines.	2008	2, 3
Chung, JY, Wan, CX, Liao, GQ	Research on Neural Network Cost Prediction Model Based on the Rough Sets Theory in ERP.	2008	2, 3
Sudzina, F, Pucihar, A, Lenart, G	ERP Systems selection criteria: a comparative study of Slovak and Slovenian SMEs.	2008	2, 3
Sommestad, T, Lillieskold, J	Development of an effort estimation model: A case study on delivery projects at a leading IT provider within the electric utility industry.	2007	2, 3
Giannopoulos, N., Taratoukhine, V., Roy, R,	Integration of Enterprise Resource Planning (ERP) and Cost Estimating (CE) systems - The challenges	2001	2, 3
Kusters, R.J., Genuchten van, M. J. I. M., Heemstra, F. J.	Are software cost-estimation models accurate?	1990	2, 3
Netjes, M., Mansar, S. L., Reijers, H. A., Aalst van der, W. M. P.	An evolutionary approach for business process design towards an intelligent system	2007	2, 3
Kusters, R.J., Heemstra, F.J., Jonker, A.	ERP Implementation Costs: A Preliminary Investigation.	2008	5
Freying, M	Estimation of the impact of enterprise resource planning project management decisions on post implementation maintenance cost: a case study using simulation modelling.	2010	2, 3

APPENDIX III: GERMAN ORIGINAL VERSION OF QUESTIONNAIRE

Druckversion

Fragebogen

1 Anfang

Herzlich Willkommen !

Ich danke Ihnen dass Sie sich dafür entschieden haben an meiner Online-Umfrage zum Thema:

Kostenstruktur und Kostentreiber bei der Auswahl, Einführung und dem Betrieb von ERP Systemen
teilzunehmen. Die Umfrage gliedert sich in die folgenden sechs Abschnitte A bis F und beinhaltet insgesamt 49 Fragen.

Abschnitt A : Unternehmensdaten
Abschnitt B : Kostenstruktur ERP Auswahl
Abschnitt C : Kostenstruktur ERP Einführung
Abschnitt D : Kostenstruktur laufender ERP Betrieb
Abschnitt E : Technische Kostentreiber
Abschnitt F : Projektbezogene Kostentreiber

Im **Abschnitt A** werden Sie gebeten einige Unternehmensdaten anzugeben. Dies dient der Klassifizierung Ihres Unternehmens.
In den **Abschnitten B bis D** werden Sie gebeten jeweils Angaben zu Ihren Aufwendungen für die Auswahl, die Einführung und die Aufwendungen
für den jährlichen Betrieb Ihres ERP-Systems anzugeben. Sie können die Aufwendungen in Euro oder Manntagen angeben.
In den **Abschnitten E und F** werden Sie gebeten technische Spezifikationen Ihres ERP-Systems anzugeben und projektbezogene Einflussgrößen in Ihrem Unternehmen zu bewerten.

Die Beantwortung nimmt insgesamt ca. 20 Minuten in Anspruch.

2 Block A Unternehmensdaten

Abschnitt A: Unternehmensdaten

Die nun folgenden Fragen zu Ihrem Unternehmen dienen dazu Ihr Unternehmen für den weiteren Verlauf der Umfrage nach Größe und Ausrichtung klassifizieren zu können.

Bitte nennen Sie...

1. Welches ERP System wurde in Ihrem Unternehmen eingeführt?
2. In welchem Jahr wurde in Ihrem Unternehmen ein ERP System eingeführt?
3. An wie vielen Standorten ist Ihr Unternehmen mit dem ERP System tätig?
4. Aus wie vielen Abteilungen besteht Ihr Unternehmen?
5. Nennen Sie bitte die Höhe des Jahresumsatzes zum Zeitpunkt der ERP Einführung:
6. In welcher Branche ist Ihr Unternehmen tätig?
7. Wie viele Mitarbeiter beschäftigt Ihr Unternehmen?

3 Block B Kostenstruktur ERP Auswahl

In den Abschnitten B, C und D werden Sie nun gebeten Angaben zu den Aufwendungen bzgl. Ihres ERP-Systems zu machen.
Dabei müssen die Aufwendungen unbedingt zwischen **Auswahl, Einführung und dem Betrieb** unterschieden werden.
Die Aufwendungen können Sie in fünf unterschiedliche Kostenarten wie **interne Personalkosten, externe Personalkosten, Hardwarekosten, ERP Softwarekosten und Lizenzkosten** aufteilen.

Falls Sie den genauen Wert nicht kennen, geben Sie bitte eine möglicht genaue Schätzung an.

Abschnitt B: Kostenstruktur ERP Auswahl

Machen Sie in diesem Abschnitt B bitte ausschließlich Angaben zu den Aufwendungen Ihrer **ERP Auswahl**.
Anders gesagt: Wieviel Geld bzw. Zeit hat Ihr Unternehmen für die Auswahl Ihres ERP Systems investiert?
Die Kostenarten für die ERP Auswahl beschränken sich auf die **internen** (Mitarbeiter) und **externen** (Berater) **Personalkosten**
Sie können den Aufwand in Euro (falls bekannt) oder in Manntagen angeben.

Bitte nennen Sie...

8. Interne Personalkosten (Mitarbeiter) für die **Auswahl** Ihres ERP Systems:

Euro Manntage

9. Externe Personalkosten (z.B. Beraterkosten) für die **Auswahl** Ihres ERP Systems:

Euro Manntage

4 Block C Kostenstruktur ERP Einführung

Abschnitt C: Kostenstruktur ERP Einführung

Machen Sie in diesem Abschnitt C bitte ausschließlich Angaben zu den Aufwendungen Ihrer **ERP Einführung**.
Der betrachtete Zeitraum ist dabei vom Zeitpunkt nach Abschluss der ERP Auswahlphase bis zum sogenannten "Echtstart" oder "Going Live".
Sollte Ihr System zum Echtstart nicht vollständig eingeführt worden sein betrachten Sie bitte den Zeitpunkt bis ca. 80 % des geplanten
Einführungsumfangs abgeschlossen waren als Ende der Einführungsphase.
Teilen Sie die in diesem Zeitrum angefallenden Kosten bitte zwischen den folgenden Kostenarten auf. Sie können den Aufwand in Euro (falls bekannt) oder wenn sinnvoll in Manntagen angeben.

Bitte nennen Sie...

10. Interne Personalkosten (inkl. **intern** durchgeführter Sonderanpasungen) bei der **Einführung** Ihres ERP Systems:

Euro ______ Manntage ______

11. Externe Personalkosten (z.B. Beraterkosten) für die **Einführung** Ihres ERP Systems.

Euro ______ Manntage ______

12. Hardwarekosten für die **Einführung** Ihres ERP Systems:

Euro ______

13. ERP Softwarekosten für die **Einführung** Ihres ERP Systems (inkl. **extern** durchgeführter Sonderanpassungen und Schulungskosten insofern diese nicht schon den externen Personalkosten zugerechnet sind).

Euro ______

14. ERP Lizenzkosten bei der **Einführung** Ihres ERP Systems:

Euro ______

5 Block D Kostenstruktur ERP Betrieb

Abschnitt D: Kostenstruktur ERP Betrieb

Machen Sie in diesem Abschnitt D bitte ausschließlich Angaben zu den **jährlichen** Aufwendungen um Ihr **ERP System** in Betrieb zu halten.
Der betrachtete Zeitraum ist dabei ab dem Abschluss der ERP Einführung bzw. nach dem sogenannten "Echtstart" oder "Going Live".
Sollte Ihr System zum Echtstart nicht vollständig eingeführt worden sein, betrachten Sie bitte die Kosten ab dem Zeitpunkt nach ca. 80 %igem Abschluss des geplanten Einführungsumfangs.
Teilen Sie bitte die jährlich anfallenden Kosten auf die folgenden Kostenarten auf. Sie können den Aufwand in Euro (falls bekannt) oder wenn sinnvoll in Manntagen angeben.

Bitte nennen Sie...

15. Jährliche **interne Personalkosten** (inkl. **intern** durchgeführter Sonderanpasungen) für den **Betrieb** Ihres ERP Systems:

Euro ______ Manntage ______

16. Jährliche **externe Personalkosten** (z.B. Beraterkosten) für den **Betrieb** Ihres ERP Systems.

Euro ______ Manntage ______

17. Jährliche **Hardwarekosten** für den **Betrieb** Ihres ERP Systems:

Euro ______

18. ERP Softwarekosten für den **Betrieb** Ihres ERP Systems (inkl. **extern** durchgeführter Sonderanpassungen und Schulungskosten insofern diese nicht schon den externen Personalkosten zugerechnet sind).

Euro ______

19. Jährliche **ERP Lizenzkosten** für den **Betrieb** Ihres ERP Systems:

Euro ______

6 Block E Technische Kostentreiber

Abschnitt E: Technische Kostentreiber

Machen Sie bitte in diesem Abschnitt technische Angaben zu Ihrem ERP-Projekt.

Bitte nennen Sie...

20. Markieren Sie bitte die verwendeten Module/Prozesse aus Ihrem **ERP System**:

☐ Einkauf	☐ Vertrieb	☐ Produktion
☐ Materialwirtschaft	☐ Customer Relationship Management (CRM)	☐ Supply Chain Management (SCM)
☐ Kalkulation	☐ Finanzbuchhaltung	☐ Feinplanung
☐ Projektmanagement	☐ Lohnbuchhaltung	☐ Betriebsdatenerfassung (BDE)
☐ Dokumentenmanagement (DMS)	☐ Personalwesen (HRM)	☐ Maschinendatenerfassung (MDE)

sonstige:

21. Anzahl der Schnittstellen zu Ihrem ERP-System:

22. Wie beurteilen Sie die Komplexität der Schnittstellen ?
(1) sehr einfach bis (10) sehr komplex

1 2 3 4 5 6 7 8 9 10

23. Anzahl der EDIs zu Ihrem ERP-System (falls vorhanden):

24. Anzahl der durchgeführten Sonderanpassungen in Ihrem ERP System:

25. Anzahl der angefertigten Auswertungen und Berichte in Ihrem ERP System:

26. Beurteilen Sie allgemein die Komplexität dieser Auswertungen und Berichte?
(1) sehr einfach bis (10) sehr komplex

1 2 3 4 5 6 7 8 9 10

27. Anzahl der Nutzer Ihres ERP Systems:

28. Anzahl unterschiedlicher Benutzergruppen:

29. Beurteilen Sie allgemein die Komplexität Ihrer Daten:
(1) sehr einfach bis (10) sehr komplex

1 2 3 4 5 6 7 8 9 10

7 Block F Projektbezogene Kostentreiber

Abschnitt F: Projektbezogene Kostentreiber

Machen Sie in diesem Abschnitt bitte projektbezogene Angaben zu Ihrem ERP Einführungsprojekt bzw. geben Sie bitte bei den Fragen 32 bis 49 eine Einschätzung zu den projektbezogenen Einflüssen bei Ihrem ERP Projekt.

Bitte nennen Sie...

30. Anzahl der **externen** Mitarbeiter (z.B. Berater) bei der ERP Einführung:

31. Anzahl der **internen** Mitarbeiter die bei der ERP Einführung maßgeblich mitgewirkt haben:

32. Wie beurteilen Sie die Stimmigkeit zwischen dem ERP System und Ihrem Unternehmen?
(1) sehr gering bis (10) sehr hoch

1 2 3 4 5 6 7 8 9 10

33. Wie beurteilen Sie die fachliche Qualität Ihres Einführungsteams?
(1) sehr gering bis (10) sehr hoch

1 2 3 4 5 6 7 8 9 10

34. Für wie "eingespielt" halten SIe Ihr ERP EInführungsteam?
(1) nicht eingespielt bis (10) sehr eingespielt

1 2 3 4 5 6 7 8 9 10

35. Wie war Ihr ERP Einführungsteam in Bezug auf Abdeckung unterschiedlicher Unternehmensbereiche und Interessengruppen zusammengesetzt?
(1) sehr unausgewogen bis (10) sehr ausgewogen

1 2 3 4 5 6 7 8 9 10

36. Wie beurteilen Sie die Verfügbarkeit des Managements Ihres Unternehmens im Bezug auf das ERP Projekt?
(1) sehr hoch bis (10) sehr gering

1 2 3 4 5 6 7 8 9 10

37. Wie beurteilen Sie die Verfügbarkeit der ERP Benutzer Ihres Unternehmens im Bezug auf das ERP Projekt?
(1) sehr gering bis (10) sehr hoch

1 2 3 4 5 6 7 8 9 10

38. Wie beurteilen Sie die fachliche Qualität der bei Ihnen eingesetzten ERP Berater
(1) sehr gering bis (10) sehr hoch

1 2 3 4 5 6 7 8 9 10

39. Wie beurteilen Sie die Erfahrung der bei Ihnen eingesetzten ERP Berater ?
(1) wenig erfahren bis (10) sehr erfahren

1 2 3 4 5 6 7 8 9 10

40. Wie beurteilen Sie die kritische Einstellung Ihrer Anwender?
(1) wenig kritisch bis (10) sehr kritisch

1 2 3 4 5 6 7 8 9 10

41. Wie beurteilen Sie die allgemein die fachliche Qualität Ihrer Anwender?
(1) sehr gering bis (10) sehr hoch

1 2 3 4 5 6 7 8 9 10

42. Wie beurteilen Sie die Miteinbeziehung der Mitarbeiter Ihres Unternehmens in das ERP Projekt ?
(1) sehr hoch bis (10) sehr gering

1 2 3 4 5 6 7 8 9 10

43. Wie beurteilen Sie die Miteinbeziehung des Managements Ihres Unternehmens in das ERP Projekt?
(1) sehr gering bis (10) sehr hoch

1 2 3 4 5 6 7 8 9 10

44. Wie beurteilen Sie die Reifheit der Geschäftsprozessen Ihres Unternehmens?
(1) sehr gering bis (10) sehr hoch

1 2 3 4 5 6 7 8 9 10

45. Wie beurteilen Sie die Komplexität der Geschäftsprozesse Ihres Unternehmens?
(1) nicht komplex bis (10) sehr komplex

1 2 3 4 5 6 7 8 9 10

46. Wie beurteilen Sie allgemein die Stabilität Ihres Unternehmens?
(1) nicht stabil bis (10) sehr stabil

1 2 3 4 5 6 7 8 9 10

47. Wie beurteilen Sie die Bereitschaft zur Veränderung in Ihrem Unternehmen?
(1) sehr gering bis (10) sehr hoch

1 2 3 4 5 6 7 8 9 10

48. Wie beurteilen Sie die Motivation des ERP Einführungsteams?
(1) sehr gering bis (10) sehr hoch

1 2 3 4 5 6 7 8 9 10

49. Wie zufrieden sind Sie derzeit mit Ihrem ERP System?
(1) sehr unzufrieden bis (10) sehr zufrieden

1 2 3 4 5 6 7 8 9 10

8 Freiwillige Angaben

Vielen Dank für Ihre Mithilfe !

Abschließend möchte ich Sie noch bitten Angaben zu Ihrer Person zu tätigen um Ihnen nach Abschluss der Studie Ihr individuelles Benchmark und die Ergebniszusammenfassung zukommen zu lassen: (Emailadresse ist ausreichend)
(Die Angaben sind selbstverständlich freiwillig)

Vorname:	
Nachname:	
EMail:	
Unternehmen	

Anmerkungen:

Hier können Sie gerne Anmerkungen oder Kommentare hinterlassen

9 Endseite

Fenster schließen

APPENDIX IV: ENGLISH TRANSLATION OF QUESTIONNAIRE

Questionnaire

1 Start

Welcome!

I would like to thank you for taking part in my online survey for the following topic:

Cost structure and cost drivers during evaluation, implementation and maintenance of ERP systems
The survey is structured into the following 6 sections A to F and includes 49 questions in total.

Section A: Company details
Section B: Cost structure ERP evaluation
Section C: Cost structure ERP implementation
Section D: Cost structure ERP maintenance
Section E: Technical cost drivers
Section F: Organisational and situational cost drivers

In **section A** you will be asked to state your company details. This serves to classify your company.
In **sections B to D** you will be asked to provide details about your company's efforts for the evaluation, implementation and annual maintenance of your ERP system. Efforts can be stated in euros or person-days.
In **sections E to F** you will be invited to provide details about the technical specifications of your ERP system and assess project-related technical, organisational and situational influences within your company.

Answering of the questionnaire will take about 20 minutes.

2 Section A - Company Details

Section A: Company Details

The following questions are necessary in order to classify your company according to enterprise size and orientation.

Please indicate…

1. Which ERP system was implemented in your company?
2. In which year was the ERP system implemented in your company?
3. At how many locations does your company operate the ERP system?
4. How many departments does your company comprise?
5. What were your company's annual sales in the year of ERP implementation?
6. In which sector is your company is active?
7. How many employees does your company have?

3 Section B - Cost Structure ERP Evaluation

In sections B, C and D you will be asked to state details regarding the ERP expenditures, differentiating between efforts during **evaluation, implementation, and maintenance**.
The ERP efforts can be divided into the following five cost types: **internal personnel costs, external personnel costs, hardware costs, ERP software costs, and licence costs**.
If you are not familiar with the exact value please state an assessed value that is as accurate as possible.

Section B: Cost Structure ERP Evaluation

In this section B, please indicate only efforts for your **ERP evaluation.**
In other words: How much money and/or time did your company invest in the evaluation of your ERP system?
The cost types for ERP evaluation is limited to the **internal personnel** (employee) **costs** and **external personnel** (consultants) **costs**.
You can enter the effort in euros (if known) or person-days.

Please indicate…

8. **Internal personnel costs** (employees) for **ERP evaluation**:
   Euros:________________________ or Person days:_______________________

9. **External personnel costs** (e.g. consultant fees) for **ERP evaluation**:
   Euros:________________________ or Person days:_______________________

4 Section C - Cost Structure ERP Implementation

In this section C, please indicate only efforts for the **ERP implementation.**
The time period to be considered is from the point in time when your ERP evaluation phase was completed until the so-called "Going-Live".
If your ERP system has not been completely implemented at the Going-Live date, please use the point in time when approx. 80% of the planned implementation scope has been completed as the end of the implementation phase.
Please divide the costs incurred in this time period between the following cost types.
You can enter the effort in euros (if known) or person-days:

Please indicate…

10. **Internal personnel costs** (incl. special **internal** adjustments) during **ERP implementation**:
    Euros:________________________ or Person-days:_______________________

11. **External personnel costs** (e.g. consultant fees) for **ERP implementation**:
    Euros:________________________ or Person-days:_______________________

12. **Hardware costs** for **ERP implementation**:
    Euros:________________________ or Person-days:_______________________

13. **ERP Software costs** for **ERP implementation** (incl. special **external** adjustments and training costs, unless those are already included in the external personnel costs):
    Euros:________________________ or Person-days:_______________________

14. **ERP Licence costs** for **ERP implementation**:
    Euros:_______________________

5 Section D - Cost Structure ERP Maintenance

In this section D, please indicate only **annual** efforts for the **ERP maintenance.**
If your ERP system has not been completely implemented at the Going-Live date, please use the point in time when approx. 80% of the planned implementation scope has been completed as the end of the implementation phase.
Please divide the annual costs incurred between the following cost types.
You can enter the effort in euros (if known) or person-days.

Please indicate…

15. Annual **internal personnel costs** (incl. special **internal** adjustments) for **ERP maintenance**:
    Euros:_______________________ or Person-days:______________________

16. Annual **external personnel costs** (e.g. consultant fees) for **ERP maintenance**:
    Euros:_______________________ or Person-days:______________________

17. Annual **hardware costs** for **ERP maintenance**:
    Euros:_______________________

18. **ERP Software costs** for **ERP maintenance** (incl. special **external** adjustments and training costs, unless those are already included in the external personnel costs).
    Euros:_______________________

19. Annual **ERP Licence costs** for **ERP maintenance**:
    Euros:_______________________

6 Section E - Technical Cost Drivers

In this section, please provide technical information about your ERP project.

Please indicate…

20. Please tick the implemented modules/processes of your **ERP system**:

□ Purchasing

□ Material Management

□ Calculation

□ Project Management

□ Document Management (DMS)

□ Sales

□ Customer Relationship Management (CRM)

□ Finance

□ Payroll Accounting

□ Human Resources (HRM)

□ Production

□ Supply Chain Management (SCM)

□ Detailed Production Planning

□ Factory Data Capture

□ Machine Data Logging

Other:____

21. No. of interfaces to the ERP system: __________________

22. How do you assess the complexity of the interfaces
(1) very simple to (10) very complex:

____

23. No. of EDIs to the ERP system (if applicable): __________________

24. No. of individual programming in the ERP system: __________________

25. No. of individual reports in the ERP system: __________________

26. How do you assess the complexity of the reports -
(1) very simple to (10) very complex:
____

27. No. of ERP users: ____________________

28. No. of different ERP user groups : ____________________

29. How do you assess the complexity of your data -
(1) very simple to (10) very complex:
____

7 Section F - Organisational and Situational Cost Drivers

In this section, please provide organisational and situational information about your ERP implementation project and/or state your assessment regarding the project-related influences of your ERP project for questions 32 to 49.

Please indicate…

30. No. of **external** personnel (e.g. consultants) for ERP implementation:
____

31. No. of **internal** personnel who have been instrumental in the ERP implementation:____________________

32. How do you assess the consistency between the ERP system and your company?
(1) very low to (10) very high: ____________________________

33. How do you assess the professional qualification of the implementation team?
(1) very low to (10) very high: ____________________________

34. How “well attuned” do you consider your implementation team to be?
(1) Very bad to (10) very good: ______________________________________

35. How well was the representation of different company areas within the ERP implementation team? **(1) Very imbalanced to (10) very balanced:** ___________________

36. How do you assess the availability of the management in your company for the ERP project? **(1) Very high to (10) very low:** _________________________

37. How do you assess the availability of the ERP users in your company for the ERP project? **(1) Very low to (10) very high:** __________________________

38. How do you assess the professional qualification of the employed ERP consultants? **(1) Very low to (10) very high:** _________________________

39. How do you assess the experience of the employed ERP consultants? **(1) Not very experienced to (10) very experienced:** _______________________

40. How do you assess the critical attitude of your users? **(1) Very uncritical to (10) very critical:** _____________________________

41. How do you assess the general professional quality of your users? **(1) Very low to (10) very high:** _________________________

42. How do you assess the involvement of your employees in the ERP project? **(1) Very high to (10) very low:** ____________________________

43. How do you assess the management involvement in the ERP project? **(1) Very low to (10) very high:** ______________________________

44. How do you assess the maturity of business processes within your company? **(1) Very low to (10) very high:** _______________________

45. How do you assess the complexity of business processes within your company? **(1) Not complex to (10) very complex:** ____________________________

46. How do you assess the general stability of your organisation? **(1) Not stable to (10) very stable:** ______________________

47. How do you assess the willingness to change within your company?
    (1) Very low to (10) very high: ______________________

48. How do you assess the motivation of the ERP implementation team?
    (1) Very low to (10) very high: ____________________________

49. How satisfied are you currently with your ERP system?
    (1) Very unsatisfied to (10) very satisfied: ________________________________

8 Voluntary Information

Thank you very much for your assistance!

At the end of this survey, I would kindly ask you to provide information about yourself so that I will be able to send you your individual benchmark and the summary of results of the study: (email address is sufficient) **(Of course, this information is voluntary.)**

Name:____________________

Surname:________________________

Email address: ________________________

Company:________________________

Further Comments:

Please use this section for remarks or comments

____

____

____

____

____

APPENDIX V: INVITATION LETTER – ORIGINAL GERMAN VERSION

Eine ganz besondere Bitte - Ich benötige Ihre Hilfe bei meiner Doktorarbeit

---

Sehr geehrte Damen und Herren,

ich möchte Sie bitten an meiner Umfrage teilzunehmen bzw. diese Email an den ERP Verantwortlichen in Ihrem Hause weiterzuleiten.

Im Rahmen meiner Dissertation führe ich eine wissenschaftliche Untersuchung zu der Kostenstruktur und den Kostentreibern bei der Auswahl, Einführung und dem Betrieb von Enterprise Resource Planning (ERP) Systemen in deutschen KMUs durch. Ziel der Arbeit ist es Kostentreiber und deren Abhängigkeiten und Auswirkungen bei Einführungsprojekten von ERP-Systemen ausfindig zu machen. Wenn Ihr Unternehmen in den letzten zehn Jahren ein ERP System eingeführt hat, möchte ich Sie herzlich bitten an meiner Umfrage teilzunehmen. Das Ausfüllen des Fragebogens nimmt ca. 20 Minuten in Anspruch und kann bequem online durchgeführt werden.

Zur Umfrage

Sämtliche von Ihnen zur Verfügung gestellten Informationen werden selbstverständlich streng vertraulich behandelt und ausschließlich im Zuge meiner Dissertation verwendet. Dabei werden in keinsterweise Unternehmensnamen oder sonstige Details Ihres Unternehmens preisgegeben.

Als Dankeschön für Ihre Mühe erhalten Sie eine Ergebniszusammenfassung und ein individuelles Benchmark Ihres ERP-Einführungsprojektes.

Einige evtl. auftretende Fragen, die ggf. vor oder während der Beantwortung der Umfrage auftreten können, werden unten im FAQ Bereich beantwortet.

Herzlichen Dank !

Benjamin Liehr

Zur Umfrage

F A Q

1. Warum sollte ich an der Umfrage teilnehmen?
2. Wer führt diese Umfrage durch?
3. Wer kann an dieser Studie teilnehmen?
4. Wie kann ich an der Befragung teilnehmen?
5. Was geschieht mit den von mir bereitgestellten Daten und den Ergebnissen der Studie?
6. Kann ich als Teilnehmer die Ergebnisse der Studie einsehen?
7. Wen kann ich im Falle von technischen Problemen oder weiteren Fragen ansprechen?
8. Welche Daten werden während der Umfrage abgefragt?
9. Was gibt es bei der Beantwortung der Umfrage zu beachten?

1. Warum sollte ich an der Umfrage teilnehmen?

Es gibt einige gute Gründe die dafür sprechen an dieser Studie teilzunehmen:

- Sie erhalten ein individuelles Benchmark zu Ihrem ERP Einführungsprojekt und eine Ergebniszusammenfassung der Studie.
- Sie haben die Möglichkeit, Ihr ERP Projekt hinsichtlich Kosten und Projekteinflüssen strukturiert zu reflektieren.
- Durch Ihre Teilnahme an der Studie ermöglichen Sie es ein besseres Verständnis über die Abhängigkeiten von ERP Projektkosten zu generieren und daraus wertvolle Erkenntnisse für die Praxis abzuleiten.
- Sie helfen mir damit sehr bei meiner Doktorarbeit.

zur Umfrage | Seitenanfang | FAQ Übersicht

2. Wer führt diese Umfrage durch?

Mein Name ist Benjamin Liehr, ich bin Jahre alt und führe im Rahmen meiner Dissertation an der Gloucestershire University, England in Kooperation mit der Fachhochschule des Mittelstandes in Bielefeld diese Studie durch.

zur Umfrage | Seitenanfang | FAQ Übersicht

3. Wer kann an dieser Studie teilnehmen?

Die Studie richtet sich hauptsächlich an Geschäftsführer, Führungskräfte oder Projektleiter bzw. Mitarbeiter die bei der Einführung eines ERP-Systems maßgeblich mitgewirkt haben bzw. über entsprechendes Know-how verfügen. Unternehmen die an dieser Umfrage teilnehmen, sollten idealerweise im weiten Sinne aus dem produzierenden Gewerbe kommen und eine Mitarbeiterzahl von min. 50 Mitarbeitern aufweisen. Auch angrenzende Industrien sind herzlich eingeladen an dieser Umfrage teilzunehmen.

zur Umfrage | Seitenanfang | FAQ Übersicht

4. Wie kann ich an der Befragung teilnehmen?

Sie haben mehrere Möglichkeiten an der Studie teilzunehmen:

- Die einfachste Variante ist an der Onlineumfrage teilzunehmen. Dazu klicken Sie einfach auf einen der Hyperlinks "zur Umfrage".
- Sie können auch folgenden Link in die Adresszeile Ihres Browsers kopieren und mit Enter bestätigen: http://ww3.unipark.de/uc/liehr_Gloucestershire_University/25f1/?code=1e859a4f31078d5a
- Sie können die Umfrage auch in Form eines schriftlichen Formulars ausfüllen. Kontaktieren sie mich dazu bitte per Mail:
- Wir können die Umfrage in Form eines persönlichen oder telefonischen Interviews durchführen. Kontaktieren Sie mich dazu bitte ebenfalls per Mail.

zur Umfrage | Seitenanfang | FAQ Übersicht

5. Was geschieht mit den von mir bereitgestellten Daten und den Ergebnissen der Studie?

Die von Ihnen breitgestellten Daten fließen in meine wissenschaftliche Arbeit ein. Auf der Basis Ihrer Datensätze werden die Daten zunächst aufbereitet und analysiert. Da es sich dabei um eine quantitative Analyse handelt wird in meiner Disseration **nicht** auf einzelne Unternehmen eingegangen.
Daher bleiben Ihre Daten nicht nur anonym sondern werden gar nicht erst einzeln aufgeführt.

zur Umfrage | Seitenanfang | FAQ Übersicht

6. Kann ich als Teilnehmer die Ergebnisse der Studie einsehen?

Sie haben nach Abschluss der Umfrage die Möglichkeit Ihre Kontaktdaten zu hinterlegen. Nach Abschluss der Studie erhalten Sie dann eine Ergebniszusammenfassung und ein individuelles Benchmark für Ihr ERP Einführungsprojekt.

zur Umfrage | Seitenanfang | FAQ Übersicht

7. Wen kann ich im Falle von technischen Problemen ansprechen?

Sollten während der Umfrage technische Probleme auftreten, möchte ich Sie bitten sich per Mail an mich zu wenden. Wir können dann eine Alternative besprechen um den Fragebogen auszufüllen.
Für weitere Fragen stehe ich Ihnen jederzeit unter folgender Adresse zur Verfügung:

zur Umfrage | Seitenanfang | FAQ Übersicht

8. Welche Daten werden während der Umfrage abgefragt?

- Zunächst werden Sie gebeten größenklassifizierende Daten zu Ihrem Unternehmen wie z.B. die Mitarbeiterzahl und den Jahresumsatz anzugeben.
- In den folgenden drei Abschnitten werden Sie aufgefordert Ihre Aufwendungen für die Auswahl, die Einführung und den jährlichen Betrieb Ihres ERP Systems anzugeben. Diese Aufwendungen müssen jeweils auf die fünf Kostenarten: interne Personalkosten, externe Personalkosten, Softwarekosten, Lizenzkosten und Hardwarekosten aufgeteilt werden.
- In den beiden darauf folgenden und gleichzeitig letzten Abschnitten ist Ihre Einschätzung zu bestimmten Kostentreibern in Form von technischen ERP-Spezifikationen und weiteren Projekteinflussgrößen gefragt. Sie werden gebeten eine Einschätzung auf einer Skala von 1-10 abzugeben.

zur Umfrage | Seitenanfang | FAQ Übersicht

9. Was gibt es bei der Beantwortung der Umfrage zu beachten?

Dieser Abschnitt "9." soll Sie beim Ausfüllen der Umfrage unterstützen. Es ist sehr empfehlenswert diesen Abschnitt vor der Start einmal aufmerksam durchzulesen und ggf. auszudrucken um ihn beim ausfüllen der Umfrage zur Hand zu haben.

Phasen des zugrungeliegenden Lebenszyklusmodells

Grundlage dieser Studie ist ein vereinfachtes Lebenszyklusmodell eines ERP Systems dass sich aus den Phasen

- Auswahl
- Einführung
- Betrieb

zusammensetzt. Häufig ist die Abgrenzung zwischen Einführung und dem Betrieb nicht eindeutig zu differenzieren. Dennoch ist es notwendig dass diese beiden Phasen unterschieden werden. Wenn Sie es nicht genau abgrenzen können nehmen Sie im Zweifel bitte eine Schätzung vor wenn 80% des geplanten Einführungsumfangs erreicht sind. Ab diesem Zeitpunkt sollte die "Einführung" dann abgeschlossen sein und die Phase des "Betriebs" beginnt. Beachten Sie bitte, dass die Kosten für die Auswahl und die Einführung nur einmal anfallen. Die Betriebskosten hingegen sind jährliche, sich widerholende, Kosten.

Abgrenzung der Kostenarten

Sie werden im Laufe der Umfrage gebeten die Aufwendungen für die Auswahl, die Einführung und den Betrieb Ihres ERP Systems jeweils auf fünf unterschiedliche Kostenarten aufzuteilen: interne Personalkosten, externe Personalkosten, Softwarekosten, Lizenzkosten, Hardwarekosten. Da eine Abgrezung manchmal nicht ganz eindeutig sein kann, sollen die folgenden Definitionen der einzelnen Kostenarten und die Nennung einiger Beispiele etwas mehr Klarheit schaffen.

Interne Personalkosten: Sämtliche interne Personalaufwendungen die Ihrem Unternehmen entstanden sind, die im Zusammenhang mit Ihrem ERP-System stehen. Dazu zählen z.B.:

- Zeitliche bzw. finanzielle Aufwendungen für den Besuch von Fachmessen für die Auswahl eines geeigneten ERP Systems.
- Aufwendungen für die Bereitstellungen von eigenem Personal wie z.B. Key-User Team, IT-Abteilung für das ERP-System, hinsichtlich Auswahl, Einführung, und Betrieb.
- Personalaufwand für das Instandhalten Ihres ERP Systems.
- Sollten Sie Personal über Zeit-und Leiharbeitsfirmen für Ihr ERP Projekt eingesetzt haben, ordnen Sie diese Aufwendungen bitte ebenfalls den internen Personalkosten zu.

Externe Personalkosten: Aufwendungen für externes Personal. Darunter fallen z.B. Beraterkosten. Beispiele sind:

- Beraterkosten für die Auswahl Ihres ERP Systems.
- ERP Beraterkosten für Workshops oder Installationen.
- Beraterkosten für Geschäftsprozessanalysen insofern dieser in direktem Zusammenhang mit Ihrem ERP System stehen.

Softwarekosten: Häufig sind die Kostenmodelle einzelner ERP Anbeiter recht unterschiedlich und lassen sich nur schwer vergleichen. Softwarekosten sollen in diesem Zusammenhang Kosten darstellen, die unmittelbar mit der ERP-Software in Verbindung stehen und nur einmalig anfallen ansonsten sind diese unte Lizenzkosten anzugeben. Unter Softwarekosten fallen z.B.:

- Einmalige Aufwendungen bei dem "Kauf" eines ERP Systems.
- Schulungskosten (insofern diese nicht bereits über externe Personalkosten berücksichtigt sind)
- Kosten für Sonderanpassungen Ihres ERP Systems insofern diese **extern** durchgeführt wurden. Sonderanpassungen die durch eigenes Personal erstellt werden sollte unter internen Personalkosten ausgewiesen werden.

Lizenzkosten: Lizenzkosten sind Softwarekosten die regelmäßig anfallen.

Hardwarekosten: Geben Sie bitte bei den Hardwarekosten die Investitionen an, die notwendig waren um Ihre IT Infrastruktur für den Betrieb Ihres ERP-Systems zu verändern bzw. durchschnittlich jährlich notwendig sind um Ihr ERP-System Ihr instandzuhalten.

Bewertungsfragen

Ein Großteil der Umfrage besteht aus Bewertungsfragen. Sie sind hier gefragt eine möglichst genaue Einschätzung zu geben. Sie haben dabei die Möglichkeit eine Frage auf Grundlage eines gegensätzlichen

Wertepaares zwischen 1 und 10 zu bewerten. Die Antwortmöglichkeiten sind äquidistant. Die Abstände der Antwortmöglichkeiten sind also als gleich zu betrachten.
Eine neutrale Positionierung läge daher theoretisch bei 5,5. Eine neutrale Bewertung ist in dieser Umfrage allerdings ausgeschlossen. Sie müssten sich daher also in so einem Falle für eine mehr zustimmende oder mehr nicht zustimmende Antwort (also zwischen 5 oder 6) entscheiden.

zur Umfrage | Seitenanfang | FAQ Übersicht

APPENDIX VI: ENGLISH TRANSLATION OF INVITATION LETTER

A special request – I would like to ask for your assistance for my doctoral thesis

Dear Sir or Madam,

I am writing to you to kindly ask you to participate in my survey or to forward this email to the person in charge of the ERP system in your company.

Within the scope of my doctoral thesis, I am conducting a scientific research regarding the cost structure and the cost drivers during the evaluation, implementation and maintenance of Enterprise Resource Planning (ERP) systems in German SMEs. The aim of my work is to determine cost drivers and their dependencies and effects on ERP implementation projects. If your company implemented an ERP system during the last ten years, I would kindly ask you to participate in my survey. The completion of the questionnaire can be done online and will take about 20 minutes.

To the survey

All of the information provided by you will, of course, be treated as strictly confidential and will only be used within the scope of my doctoral thesis. The name of your company or any other details about your company will not be disclosed.

As a small token of appreciation for your participation you will receive a results summary and an individual benchmark for your ERP implementation project.

Any questions that may arise prior to or during the survey participation are explained in the FAQ section below.

Thank you very much indeed!

Benjamin Liehr

To the survey

F A Q

1. Why should I participate in the survey?

There are some very good reasons for participating in this survey:

- You will receive an individual benchmark for your ERP implementation project and a summary of the results identified in this study.
- You will have the opportunity to reflect on costs and project influences in your own ERP project.
- Participation in this study will enable you to gain a better understanding of the dependencies of ERP project costs and obtain valuable practical information.
- You are providing valuable assistance to me for my doctoral thesis.

2. Who is conducting this survey?

My name is Benjamin Liehr. I am years old and conducting this survey for my doctor-al thesis at the University of Gloucestershire in England in cooperation with the Fachhochschule des Mittelstandes (University of Applied Sciences with a special focus on SMEs), Bielefeld.

3. Who can participate in this survey?

The target group for this survey is mainly managing directors, executives or project managers and/or employees who were instrumental in the implementation of an ERP system and/or who have the respective know-how. Companies participating in this survey should ideally, in a broad sense, be in the manufacturing sector and have at least 50 employees. Related sectors are also welcome to participate in this survey.

4. How can I participate in the survey?

You have several options for participating in this survey:

- The easiest way to participate is to answer the online questionnaire. To do so, just click on one of the hyperlinks “To the survey”.
  Alternatively, you can copy the following link into the address bar of your browser and confirm with “Enter”:
  http://ww3.unipark.de/uc/liehr_Gloucestershire_University/25f1/?code=1e859a4f31078d5a
- You can also participate in the survey by filling out a printed form. To do so, please contact me at my email address:
- We can conduct this survey by way of an in-person or telephone interview. For this option, please also contact me by email.

5. What happens with the data provided by me and with the results of the study?

The data provided by you will be incorporated into my scientific paper. Initially, the data will be processed and analysed on the basis of your data sets. Since this is a quantitative analysis, my doctoral thesis will **not** focus on individual companies. This means that your data remains anonymous and will not even be individually mentioned.

6. Can I, as a participant, get access to the results?

After completing the survey, you will have the opportunity to indicate your personal data. After completing my survey, I will send you a report on the results and an individual benchmark for your ERP implementation project.

7. Whom can I contact in case of a technical problem?

In case of a technical problem, I would ask you to contact me by email so we can discuss an alternative way of completing the questionnaire.

I will be available to answer any further questions you may have. My email address is

8. Which data is required for the survey?

- The first part is answering size-classifying data about your company, like the number of employees or the annual turnover.
- The next three sections are about the expenses/costs for the selection, the implementation and the annual maintenance of your ERP system. The costs are divided into five cost types: internal personnel, external personnel, software, licence, and hardware.
- The two subsequent and also last sections are about giving your assessment of certain cost drivers in the form of technical ERP specifications and other project parameters. You will be asked to give your assessment by using a 1 to10 scale.

9. What do I need to consider when completing the questionnaire?

This section “9” should support you in the completion of the questionnaire. It is strongly recommended that you read this section carefully before starting the survey. It might be helpful to print it out in order to have it on hand when completing the survey.

Phases of underlying Lifecycle Model

The basis of my study is a simplified lifecycle model of an ERP system consisting of the following three phases:

- Evaluation
- Implementation
- Maintenance

Often, a clear distinction between "implementation" and "maintenance" is hard to make. However, it is necessary to differentiate between these two phases. If you cannot make an exact differentiation, please use an estimate if 80% of the planned scope of implementation is completed. From this point in time, the "implementation" should be completed and the "maintenance" phase should start. Please bear in mind that the costs of evaluation and implementation are only incurred once. The maintenance costs are recurring annual costs.

Distinction between cost types

In the course of the survey, you will be asked to divide your expenses/costs for the evaluation, implementation and maintenance of your ERP system into five different cost types: internal personnel costs, external personnel costs, software, licence, and hardware costs. Since it is sometimes difficult to make a distinction, the following definitions and examples serve to clarify the meaning of the individual cost types:

Internal personnel costs: All internal personnel costs incurred by your company and associated with the ERP system. These include, e.g.:

- Time and/or financial expenditures for visiting ERP fairs to select a suitable ERP system
- Expenses for the allocation of internal personnel, such as key user team and IT department, with regard to the evaluation, implementation, and the maintenance of an ERP system
- Internal personnel efforts for the maintenance of your ERP system
- If you have used personnel for your ERP project through temporary employment agencies, please include these expenses in the internal personnel costs.

External personnel costs: Expenses for external personnel. These include e.g. consultant fees. Examples:

- Consultant fees for the evaluation of your ERP system
- ERP consultant fees for installation work or workshops
- Consultant fees for business process analyses, provided these are directly associated with the ERP system

Software costs: Price policies of individual ERP vendors offer differ considerably, which makes a comparison difficult. Fundamentally, ERP software costs are defined in this context as costs which are directly related to the ERP system and occur only once, otherwise they should be included in licence costs. Software costs include e.g.:

- One-time expenses for the "purchase" of an ERP system.
- Training costs
  (insofar as these efforts are not already considered in the external personnel costs)
- Costs for tailoring/reprogramming of your ERP system, provided this programming work was done **externally**. Costs for programming work done in-house by the company's own personnel should be included in internal personnel costs.

Licence costs: Licence costs are software costs incurred on a regular basis.

Hardware costs: Hardware costs are the investment efforts necessary to modify your IT infrastructure for the operation of an ERP system and/or necessary annually on average to maintain your ERP system.

Assessment questions

The biggest part of the survey consists of assessment questions in which you are asked to give as precise an assessment as possible. You will have the option to assess a question based on an opposing value pair between 1 and 10. The reply options are equidistant, meaning the distances of the reply options should be considered as equal.

Theoretically, a neutral position would thus be at 5.5. However, a neutral assessment is excluded in this survey. Therefore, in a case like that, you would have to decide on a more agreeing or more disagreeing answer (i.e. between 5 or 6).